PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, 16/18, C12N 15/62

A1

(11) International Publication Number: **WO 99/03994**

(43) International Publication Date: 28 January 1999 (28.01.99)

(21) International Application Number: PCT/US98/14545

(22) International Filing Date: 16 July 1998 (16.07.98)

(30) Priority Data:

60/053,153	18 July 1997 (18.07.97)	US
60/053,244	18 July 1997 (18.07.97)	US
60/055,804	15 August 1997 (15.08.97)	US
60/055,853	15 August 1997 (15.08.97)	US

(71) Applicant: SCHERING CORPORATION [US/US]; 2000 Galloping Hill Road, Kenilworth, NJ 07033 (US).

(72) Inventor: JOHNSON, James, A.; 205 Mary Alice Drive, Los Gatos, CA 95032 (US).

(74) Agents: MCLAUGHLIN, Jaye, P. et al.; Schering-Plough Corporation, Patent Dept., K-6-1 1990, 2000 Galloping Hill Road, Kenilworth, NJ 07033-0530 (US).

(81) Designated States: AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CZ, EE, GE, HR, HU, ID, IL, IS, JP, KG, KR, KZ, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UZ, VN, YU, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: SUPPRESSORS OF CYTOKINE SIGNALING SOCS16; RELATED REAGENTS

(57) Abstract

Purified genes encoding intracellular regulatory molecules from a human, reagents related thereto including purified proteins, specific antibodies, and nucleic acids encoding these molecules are provided. Methods of using said reagents and diagnostic kits are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

SUPPRESSORS OF CYTOKINE SIGNALING SOCS16; RELATED REAGENTS

This filing is a PCT filing claiming priority to provisional U.S. Patent Applications USSN 60/055,853, filed August 15, 1997; USSN 60/053,244, filed July 18, 1997; USSN 60/055,804, filed August 15, 1997; and USSN 60/053,153, filed July 18, 1997; each of which is incorporated herein by reference.

FIELD OF THE INVENTION

The present invention pertains to compositions related to proteins which function, e.g., in suppressing intracellular signaling pathways, e.g., cytokine signaling. In particular, it provides purified genes, proteins, antibodies, and related reagents useful, e.g., to regulate growth hormone-like or cytokine-regulated intracellular processes, including transcription or genes in various cell types, including immune cells.

BACKGROUND OF THE INVENTION

Recombinant DNA technology refers generally to the technique of integrating genetic information from a donor source into vectors for subsequent processing, such as through introduction into a host, whereby the transferred genetic information is copied and/or expressed in the new environment. Commonly, the genetic information exists in the form of complementary DNA (cDNA) derived from messenger RNA (mRNA) coding for a desired protein product. The carrier is frequently a plasmid having the capacity to incorporate cDNA for later replication in a host and, in some cases, actually to control expression of the cDNA and thereby direct synthesis of the encoded product in the host.

For some time, it has been known that the mammalian immune response is based on a series of complex cellular interactions, called the "immune network". Recent

research has provided new insights into the inner workings of this network. While it remains clear that much of the response does, in fact, revolve around the network-like interactions of lymphocytes, macrophages, granulocytes, and other cells, immunologists now generally hold the opinion that soluble proteins, known as lymphokines, cytokines, or monokines, play a critical role in controlling these cellular interactions. Thus, there is considerable interest in the isolation, characterization, and mechanisms of action of cell modulatory factors, an understanding of which will lead to significant advancements in the diagnosis and therapy of numerous medical abnormalities, e.g., immune system disorders. Some of these factors are hematopoietic growth factors, e.g., granulocyte colony stimulating factor (G-CSF), and others are regulatory molecules. See, e.g., Thomson (1994; ed.) The Cytokine Handbook (2d ed.) Academic Press, San Diego; Metcalf and Nicola (1995) The Hematopoietic Colony Stimulating Factors Cambridge University Press; and Aggarwal and Gutterman (1991) Human Cytokines Blackwell Pub.

Lymphokines apparently mediate cellular activities in a variety of ways. They have been shown to support the proliferation, growth, and differentiation of, e.g., pluripotential hematopoietic stem cells into vast numbers of progenitors comprising diverse cellular lineages making up a complex immune system. Proper and balanced interactions between cellular components are necessary for a healthy developmental or immune response. The different cellular lineages often respond in a different manner when lymphokines are administered in conjunction with other agents.

In the immune system, many of the effects of known cytokines on gene transcription are known to be mediated by cytokine inducible DNA binding proteins. See, e.g., Paul (ed. 1994) Fundamental Immunology, 3rd ed., Raven Press, New York, NY. The mechanisms of signal

transduction have been an area of active recent study, and involve protein phosphorylation and dephosphorylation with, e.g., the Janus kinases (JAKs) and Signal Transducers and Activators of Transcription (Stats). See, e.g., Ihle (1996) Cell 84:331-334; ;Ivashkiv (1995) Immunity 3:1-4; and Ihle and Kerr (1995) Trends in Genetics 11:69-74.

The lack of knowledge regarding the mechanisms of signaling involved in the regulation of cell cycle or transcriptional elements has hampered the ability of medical science to specifically regulate cell division or cellular responses, including immune responses. The present invention provides compositions which will be important in such regulation.

SUMMARY OF THE INVENTION

The present invention is based in part upon the discovery of intracellular regulatory molecules which can block signal transduction, e.g., through growth factor- or cytokine-receptor superfamily signaling mechanisms. These proteins exhibit a structural feature designated a SOCS box. See Hilton, et al. (1998) Proc. Nat'l Acad. Sci. USA 95:114-119. Moreover, the SOCS3 protein can block the IL-2 induced signaling via the STAT5, establishing function of the SOCS proteins as suppressors of cytokine signaling.

The invention provides a substantially pure or recombinant SOCS16 protein or peptide exhibiting at least about 85% sequence identity over a length of at least about 12 amino acids to SEQ ID NO: 2 or 4; a natural sequence SOCS16 of SEQ ID NO: 2 or 4; or a fusion protein comprising SOCS16 sequence. In preferred embodiments, the substantially pure or isolated protein comprises a segment exhibiting sequence identity to a corresponding portion of a SOCS16 wherein: the homology is at least about 90% identity and the portion is at least about 9 amino acids; the homology is at least about 80% identity

and the portion is at least about 17 amino acids; or the homology is at least about 70% identity and the portion is at least about 25 amino acids. In other embodiments, the: SOCS16 comprises a mature sequence of SEQ ID NO: 2 or 4; protein or peptide: is from a warm blooded animal selected from a mammal, including a primate; comprises at least one polypeptide segment of SEQ ID NO: 2 or 4; exhibits a plurality of portions exhibiting the identity; is a natural allelic variant of SOCS16; has a length at least about 30 amino acids; exhibits at least two non-overlapping epitopes which are specific for a mammalian SOCS16; exhibits a sequence identity at least about 90% over a length of at least about 20 amino acids to a primate SOCS16; exhibits at least two non-overlapping epitopes which are specific for a primate SOCS16; exhibits a sequence identity at least about 90% over a length of at least about 20 amino acids to a primate SOCS16; is glycosylated; is a synthetic polypeptide; is attached to a solid substrate; is conjugated to another chemical moiety; is a 5-fold or less substitution from natural sequence; or is a deletion or insertion variant from a natural sequence. Various preferred embodiments include a composition comprising: a sterile SOCS16 protein or peptide; or the SOCS16 protein or peptide and a carrier, wherein the carrier is: an aqueous compound, including water, saline, and/or buffer; and/or formulated for oral, rectal, nasal, topical, or parenteral administration. The invention further provides a fusion protein, comprising: mature protein comprising sequence of SEQ ID NO: 2 or 4; a detection or purification tag, including a FLAG, His6, or Ig sequence; or sequence of another SOCS protein.

These reagents also make available a kit comprising such a protein or polypeptide, and: a compartment comprising the protein or polypeptide; and/or instructions for use or disposal of reagents in the kit.

Providing an antigen, the invention further provides a binding compound comprising an antigen binding portion from an antibody, which specifically binds to a natural SOCS16 protein, wherein: the protein is a primate protein; the binding compound is an Fv, Fab, or Fab2 fragment; the binding compound is conjugated to another chemical moiety; or the antibody: is raised against a peptide sequence of a mature polypeptide comprising sequence of SEQ ID NO: 2 or 4+ ; is raised against a mature SOCS16; is raised to a purified SOCS16; is immunoselected; is a polyclonal antibody; binds to a denatured SOCS16; exhibits a Kd to antigen of at least 30 μM; is attached to a solid substrate, including a bead or plastic membrane; is in a sterile composition; or is detectably labeled, including a radioactive or fluorescent label. Preferred kits include those containing the binding compound, and: a compartment comprising the binding compound; and/or instructions for use or disposal of reagents in the kit. Many of the kits will be used for making a qualitative or quantitative analysis.

Other preferred compositions will be those comprising: a sterile binding compound, or the binding compound and a carrier, wherein the carrier is: an aqueous compound, including water, saline, and/or buffer; and/or formulated for oral, rectal, nasal, topical, or parenteral administration.

The present invention further provides an isolated or recombinant nucleic acid encoding a protein or peptide or fusion protein described above, wherein: the SOCS family protein is from a mammal, including a primate; or the nucleic acid: encodes an antigenic peptide sequence of SEQ ID NO: 2 OR 4; encodes a plurality of antigenic peptide sequences of SEQ ID NO: 2 AOR 4; exhibits at least about 80% identity to a natural cDNA encoding the segment; is an expression vector; further comprises an origin of replication; is from a natural source;

comprises a detectable label; comprises synthetic nucleotide sequence; is less than 6 kb, preferably less than 3 kb; is from a mammal, including a primate; comprises a natural full length coding sequence; is a hybridization probe for a gene encoding the SOCS family protein; or is a PCR primer, PCR product, or mutagenesis primer. In certain embodiments, the invention provides a cell or tissue comprising such a recombinant nucleic acid. Preferred cell s include: a prokaryotic cell; a eukaryotic cell; a bacterial cell; a yeast cell; an insect cell; a mammalian cell; a mouse cell; a primate cell; or a human cell.

Other kit embodiments include a kit comprising the described nucleic acid, and: a compartment comprising the nucleic acid; a compartment further comprising a SOCS16 protein or polypeptide; and/or instructions for use or disposal of reagents in the kit. In many versions, the kit is capable of making a qualitative or quantitative analysis.

Other nucleic acid embodiments include those which: hybridize under wash conditions of 40° C and less than 500 mM salt to SEQ ID NO: 1 or 3; or exhibits at least about 85% identity over a stretch of at least about 30 nucleotides to a primate SOCS16. In other embodiments: the wash conditions are at 45° C and/or 300 mM salt; 55° C and/or 150 mM salt; the identity is at least 90% and/or the stretch is at least 55 nucleotides; and/or at least 95% and/or the stretch is at least 75 nucleotides.

In other embodiments, the invention provides a method of modulating physiology or development of a cell or tissue culture cells comprising introducing into such cell an agonist or antagonist of a SOCS16.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. General

The proliferation, differentiation, and physiological responses of many cell lineages are regulated by secreted proteins, e.g., cytokines. These molecules often exert their biological effects through binding to cell surface receptors that are associated with one or more members of the Janus Kinase (Jak) family of cytoplasmic tyrosine kinases. For example, cytokine induced receptor dimerization leads to the activation of JAKs, rapid tyrosine phosphorylation of cytoplasmic domains, and subsequent recruitment of various signaling proteins to the receptor complex, including members of the STAT family of transcription factors. The JAK and STAT proteins are enzymes which act to transduce a signal from the cell surface to the nucleus, thereby serving as the pathway to signal the cell to respond physiologically to an external signal. These pathways have been shown to involve certain protein phosphorylation or dephosphorylation steps, thereby leading to response or lack of response by the cell. See, e.g., Ihle (1996) Cell 84:331-334; Ivashkiv (1995) Immunity 3:1-4; Ihle, et al. (1995) Ann. Rev. Immunol. 13:369-398; Ihle and Kerr (1995) Trends in Genetics 11:69-74; and Darnell, et al. (1994) Science 264:1415-1421.

A number of novel genes have been identified from mouse or humans which appear to inhibit STAT function. See, e.g., Yoshimura, et al. (1995) EMBO J. 14:2816-2826; Matsumoto, et al. (1997) Blood 89:3148-3154; Starr, et al. (1997) Nature 387:917-921; Endo, et al. (1997) Nature 387:921-924; and Naka, et al. Nature 387:924-929. The present invention provides additional genes with sequence related to those, designated Suppressors Of Cytokine Signaling, SOCS16. A Human SOCS16 CDNA fragment and the predicted amino acid translation of the open reading frame running from nucleotides 132 through 707 are provided in SEQ ID NO: 1 and 2. The translation exhibits significant

matching and similarity to other identified SOCS family members.

Additional refined primate, e.g., human sequence is provided in SEQ ID NO: 3 and 4. Nucleotide may be C or T at positions: 42, 50, 69, and 1147; nucleotide may be A or T at positions: 45, 49, 62, 66, 68, and 1152; nucleotide may be G or C at positions: 47, 52, and 1045; nucleotide may be A or C at positions: 48, 51, 54, 65, 67, 70, and 76; nucleotide may be G or T at position 64; nucleotide may be A or G at positions: 75, and 534; and nucleotide may be A, T, C, or G at positions: 1164, 1194, 1237, 1353, and 1408 (see SEQ ID NO: 6).

SOCS proteins are a family of proteins ranging from approximately 30-60 Kd which inhibit JAK kinase activity. The amino portion of SOCS proteins contain an SH2 binding motif and the carboxy portion of the molecule contains a SOCS box motif which may play a role in dimerization of SOCS proteins. The WSD are closely related in sequence.

SOCS3 expression is induced by IL-2 and can be detected by approximately 1 hour after IL-2 activation. Subsequently, SOCS expression is decreased relatively rapidly (e.g., approximately 8 hrs after activation). Western blots show that SOCS3 interacts with IL-2 receptor and JAK1 following IL-2 stimulation.

II. Definitions

The term "binding composition" refers to molecules that bind with specificity to SOCS16 protein, e.g., in an antibody-antigen interaction. However, other compounds, e.g., binding proteins, may also specifically associate with SOCS16 proteins in contrast to other molecules. Typically, the association will be in a natural physiologically relevant protein-protein interaction, either covalent or non-covalent, and may include members of a multiprotein complex, including carrier compounds or

dimerization partners. The molecule may be a polymer, or chemical reagent. A functional analog may be a protein with structural modifications, or may be a wholly unrelated molecule, e.g., which has a molecular shape which interacts with the appropriate protein binding determinants. The proteins may serve as agonists or antagonists of the binding partner, see, e.g., Goodman, et al. (eds.) (1990) Goodman & Gilman's: The Pharmacological Bases of Therapeutics (8th ed.) Pergamon Press, Tarrytown, N.Y.

The term "binding agent:SOCS16 protein complex", as used herein, refers to a complex of a binding agent and a SOCS16 protein that is formed by specific binding of the binding agent to the respective SOCS16 protein. Specific binding of the binding agent means that the binding agent has a specific binding site that recognizes a site on the SOCS16 protein. For example, antibodies raised to a SOCS16 protein and recognizing an epitope on the SOCS16 protein are capable of forming a binding agent:SOCS16 protein complex by specific binding. Typically, the formation of a binding agent:SOCS16 protein complex allows the measurement of SOCS16 protein in a mixture of other proteins and biologics. The term "antibody:SOCS16 protein complex" refers to an embodiment in which the binding agent, e.g., is an antibody. The antibody may be monoclonal, polyclonal, or a binding fragment of an antibody, e.g., an Fv, Fab, or F(ab)2 fragment. The antibody will preferably be a polyclonal antibody for cross-reactivity purposes.

"Homologous" nucleic acid sequences, when compared, exhibit significant similarity, or identity. The standards for homology in nucleic acids are either measures for homology generally used in the art by sequence comparison and/or phylogenetic relationship, or based upon hybridization conditions. Hybridization conditions are described in greater detail below.

An "isolated" nucleic acid is a nucleic acid, e.g., an RNA, DNA, or a mixed polymer, which is substantially separated from other biologic components which naturally accompany a native sequence, e.g., proteins and flanking genomic sequences from the originating species. The term embraces a nucleic acid sequence which has been removed from its naturally occurring environment, and includes recombinant or cloned DNA isolates and chemically synthesized analogs, or analogs biologically synthesized by heterologous systems. A substantially pure molecule includes isolated forms of the molecule. An isolated nucleic acid will usually contain homogeneous nucleic acid molecules, but will, in some embodiments, contain nucleic acids with minor sequence heterogeneity. This heterogeneity is typically found at the polymer ends or portions not critical to a desired biological function or activity.

As used herein, the terms "SOCS16" protein shall encompass, when used in a protein context, a protein having amino acid sequences shown in SEQ ID NO: 2 or 4 or a significant fragment of such a protein, preferably a natural embodiment. The invention also embraces a polypeptide which exhibits similar structure to human SOCS16 protein, e.g., which interacts with SOCS16 protein specific binding components. These binding components, e.g., antibodies, typically bind to a SOCS16 protein, respectively, with high affinity, e.g., at least about 100 nM, usually better than about 30 nM, preferably better than about 10 nM, and more preferably at better than about 3 nM.

The term "polypeptide" or "protein" as used herein includes a significant fragment or segment of a SOCS16 protein, and encompasses a stretch of amino acid residues of at least about 8 amino acids, generally at least 10 amino acids, more generally at least 12 amino acids, often at least 14 amino acids, more often at least 16 amino acids, typically at least 18 amino acids, more

typically at least 20 amino acids, usually at least 22 amino acids, more usually at least 24 amino acids, preferably at least 26 amino acids, more preferably at least 28 amino acids, and, in particularly preferred embodiments, at least about 30 or more amino acids, e.g., 35, 40, 45, 50, 60, 70, 80, etc. The invention encompasses proteins comprising a plurality of said segments. Features of one of the different genes should not be taken to limit those of another of the genes.

A "recombinant" nucleic acid is defined either by its method of production or its structure. In reference to its method of production, e.g., a product made by a process, the process is use of recombinant nucleic acid techniques, e.g., involving human intervention in the nucleotide sequence, typically selection or production. Alternatively, it can be a nucleic acid made by generating a sequence comprising fusion of two fragments which are not naturally contiguous to each other, but is meant to exclude products of nature, e.g., naturally occurring mutants. Thus, for example, products made by transforming cells with any non-naturally occurring vector is encompassed, as are nucleic acids comprising sequence derived using any synthetic oligonucleotide process. Such is often done to replace a codon with a redundant codon encoding the same or a conservative amino acid, while typically introducing or removing a sequence recognition site. Alternatively, it is performed to join together nucleic acid segments of desired functions to generate a single genetic entity comprising a desired combination of functions not found in the commonly available natural forms. Restriction enzyme recognition sites are often the target of such artificial manipulations, but other site specific targets, e.g., promoters, DNA replication sites, regulation sequences, control sequences, or other useful features may be incorporated by design. A similar concept is intended for a recombinant, e.g., fusion, polypeptide.

Specifically included are synthetic nucleic acids which, by genetic code redundancy, encode polypeptides similar to fragments of these antigens, and fusions of sequences from various different species variants.

"Solubility" is reflected by sedimentation measured in Svedberg units, which are a measure of the sedimentation velocity of a molecule under particular conditions. The determination of the sedimentation velocity was classically performed in an analytical ultracentrifuge, but is typically now performed in a standard ultracentrifuge. See, Freifelder (1982) Physical Biochemistry (2d ed.) W.H. Freeman & Co., San Francisco, CA; and Cantor and Schimmel (1980) Biophysical Chemistry parts 1-3, W.H. Freeman & Co., San Francisco, CA. As a crude determination, a sample containing a putatively soluble polypeptide is spun in a standard full sized ultracentrifuge at about 50K rpm for about 10 minutes, and soluble molecules will remain in the supernatant. A soluble particle or polypeptide will typically be less than about 30S, more typically less than about 15S, usually less than about 10S, more usually less than about 6S, and, in particular embodiments, preferably less than about 4S, and more preferably less than about 3S. Solubility of a polypeptide or fragment depends upon the environment and the polypeptide. Many parameters affect polypeptide solubility, including temperature, electrolyte environment, size and molecular characteristics of the polypeptide, and nature of the solvent. Typically, the temperature at which the polypeptide is used ranges from about 4° C to about 65° C. Usually the temperature at use is greater than about 18° C and more usually greater than about 22° C. For diagnostic purposes, the temperature will usually be about room temperature or warmer, but less than the denaturation temperature of components in the assay. For therapeutic purposes, the temperature will usually be body temperature, typically about 37° C for humans,

though under certain situations the temperature may be raised or lowered in situ or in vitro.

The size and structure of the polypeptide should generally be in a substantially stable state, and usually not in a denatured state. The polypeptide may be associated with other polypeptides in a quaternary structure, e.g., to confer solubility, or associated with lipids or detergents in a manner which approximates natural lipid bilayer interactions.

The solvent will usually be a biologically compatible buffer, of a type used for preservation of biological activities, and will usually approximate a physiological solvent. Usually the solvent will have a neutral pH, typically between about 5 and 10, and preferably about 7.5. On some occasions, a detergent will be added, typically a mild non-denaturing one, e.g., CHS (cholesteryl hemisuccinate) or CHAPS (3-[3-cholamidopropyl)dimethyl-ammonio]-1-propane sulfonate), or a low enough concentration as to avoid significant disruption of structural or physiological properties of the protein.

"Substantially pure" in a protein context typically means that the protein is isolated from other contaminating proteins, nucleic acids, and other biologicals derived from the original source organism. Purity, or "isolation" may be assayed by standard methods, and will ordinarily be at least about 50% pure, more ordinarily at least about 60% pure, generally at least about 70% pure, more generally at least about 80% pure, often at least about 85% pure, more often at least about 90% pure, preferably at least about 95% pure, more preferably at least about 98% pure, and in most preferred embodiments, at least 99% pure. Similar concepts apply, e.g., to antibodies or nucleic acids.

"Substantial similarity" in the nucleic acid sequence comparison context means either that the segments, or their complementary strands, when compared, are identical

when optimally aligned, with appropriate nucleotide insertions or deletions, in at least about 50% of the nucleotides, generally at least 56%, more generally at least 59%, ordinarily at least 62%, more ordinarily at least 65%, often at least 68%, more often at least 71%, typically at least 74%, more typically at least 77%, usually at least 80%, more usually at least about 85%, preferably at least about 90%, more preferably at least about 95 to 98% or more, and in particular embodiments, as high at about 99% or more of the nucleotides. Alternatively, substantial similarity exists when the segments will hybridize under selective hybridization conditions, to a strand, or its complement, typically using a sequence derived from SEQ ID NO: 1 or 3. Typically, selective hybridization will occur when there is at least about 55% similarity over a stretch of at least about 30 nucleotides, preferably at least about 65% over a stretch of at least about 25 nucleotides, more preferably at least about 75%, and most preferably at least about 90% over about 20 nucleotides. See Kanehisa (1984) Nuc. Acids Res. 12:203-213. The length of similarity comparison, as described, may be over longer stretches, and in certain embodiments will be over a stretch of at least about 17 nucleotides, usually at least about 20 nucleotides, more usually at least about 24 nucleotides, typically at least about 28 nucleotides, more typically at least about 40 nucleotides, preferably at least about 50 nucleotides, and more preferably at least about 75 to 100 or more nucleotides, e.g., 150, 200, etc.

"Stringent conditions", in referring to homology or substantial similarity in the hybridization context, will be stringent combined conditions of salt, temperature, organic solvents, and other parameters, typically those controlled in hybridization reactions. The combination of parameters is more important than the measure of any single parameter. See, e.g., Wetmur and Davidson (1968)

J. Mol. Biol. 31:349-370. A nucleic acid probe which binds to a target nucleic acid under stringent conditions is specific for said target nucleic acid. Such a probe is typically more than 11 nucleotides in length, and is sufficiently identical or complementary to a target nucleic acid over the region specified by the sequence of the probe to bind the target under stringent hybridization conditions.

SOCS16 protein from other mammalian species can be cloned and isolated by cross-species hybridization of closely related species. See, e.g., below. Similarity may be relatively low between distantly related species, and thus hybridization of relatively closely related species is advisable. Alternatively, preparation of an antibody preparation which exhibits less species specificity may be useful in expression cloning approaches.

The phrase "specifically binds to an antibody" or "specifically immunoreactive with", when referring to a protein or peptide, refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other biological components. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein and do not significantly bind other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, antibodies raised to the human protein immunogen with the amino acid sequence depicted in SEQ ID NO: 2 or 4 can be selected to obtain antibodies specifically immunoreactive with SOCS16 proteins and not with other proteins. These antibodies recognize proteins highly similar to the homologous SOCS16 protein.

III. Nucleic Acids

Human SOCS16 protein is each exemplary of a larger class of structurally and functionally related proteins. These soluble proteins will serve to transmit signals between different cell types. The preferred embodiments, as disclosed, will be useful in standard procedures to isolate genes from different individuals or other species, e.g., warm blooded animals, such as birds and mammals. Cross hybridization will allow isolation of related genes encoding proteins from individuals, strains, or species. A number of different approaches are available to successfully isolate a suitable nucleic acid clone based upon the information provided herein. Southern blot hybridization studies can qualitatively determine the presence of homologous genes in human, monkey, rat, mouse, dog, cow, and rabbit genomes under specific hybridization conditions.

Complementary sequences will also be used as probes or primers. Based upon identification of the likely amino terminus, other peptides should be particularly useful, e.g., coupled with anchored vector or poly-A complementary PCR techniques or with complementary DNA of other peptides.

Techniques for nucleic acid manipulation of genes encoding SOCS16 proteins, such as subcloning nucleic acid sequences encoding polypeptides into expression vectors, labeling probes, DNA hybridization, and the like are described generally in Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual (2nd ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY, which is incorporated herein by reference. This manual is hereinafter referred to as "Sambrook, et al."

There are various methods of isolating DNA sequences encoding SOCS16 proteins. For example, DNA is isolated from a genomic or cDNA library using labeled oligonucleotide probes having sequences identical or complementary to the sequences disclosed herein. Full-length probes may be used, or oligonucleotide probes may

be generated by comparison of the sequences disclosed. Such probes can be used directly in hybridization assays to isolate DNA encoding SOCS16 proteins, or probes can be designed for use in amplification techniques such as PCR, for the isolation of DNA encoding SOCS16 proteins.

To prepare a cDNA library, mRNA is isolated from cells which expresses a SOCS16 protein. cDNA is prepared from the mRNA and ligated into a recombinant vector. The vector is transfected into a recombinant host for propagation, screening, and cloning. Methods for making and screening cDNA libraries are well known. See Gubler and Hoffman (1983) Gene 25:263-269 and Sambrook, et al.

For a genomic library, the DNA can be extracted from tissue and either mechanically sheared or enzymatically digested to yield fragments of about 12-20kb. The fragments are then separated by gradient centrifugation and cloned in bacteriophage lambda vectors. These vectors and phage are packaged in vitro, as described in Sambrook, et al. Recombinant phage are analyzed by plaque hybridization as described in Benton and Davis (1977) Science 196:180-182. Colony hybridization is carried out as generally described in e.g., Grunstein, et al. (1975) Proc. Natl. Acad. Sci. USA. 72:3961-3965.

DNA encoding a SOCS16 protein can be identified in either cDNA or genomic libraries by its ability to hybridize with the nucleic acid probes described herein, e.g., in colony or plaque hybridization assays. The corresponding DNA regions are isolated by standard methods familiar to those of skill in the art. See, e.g., Sambrook, et al.

Various methods of amplifying target sequences, such as the polymerase chain reaction, can also be used to prepare DNA encoding SOCS16 proteins. Polymerase chain reaction (PCR) technology is used to amplify such nucleic acid sequences directly from mRNA, from cDNA, and from genomic libraries or cDNA libraries. The isolated

sequences encoding SOCS16 proteins may also be used as templates for PCR amplification.

Typically, in PCR techniques, oligonucleotide primers complementary to two 5' regions in the DNA region to be amplified are synthesized. The polymerase chain reaction is then carried out using the two primers. See Innis, et al. (eds.) (1990) PCR Protocols: A Guide to Methods and Applications Academic Press, San Diego, CA. Primers can be selected to amplify the entire regions encoding a full-length SOCS16 protein or to amplify smaller DNA segments as desired. Once such regions are PCR-amplified, they can be sequenced and oligonucleotide probes can be prepared from sequence obtained using standard techniques. These probes can then be used to isolate DNA's encoding SOCS16 proteins.

Oligonucleotides for use as probes are usually chemically synthesized according to the solid phase phosphoramidite triester method first described by Beaucage and Carruthers (1983) Tetrahedron Lett. 22(20):1859-1862, or using an automated synthesizer, as described in Needham-VanDevanter, et al. (1984) Nucleic Acids Res. 12:6159-6168. Purification of oligonucleotides is performed e.g., by native acrylamide gel electrophoresis or by anion-exchange HPLC as described in Pearson and Regnier (1983) J. Chrom. 255:137-149. The sequence of the synthetic oligonucleotide can be verified using, e.g., the chemical degradation method of Maxam, A.M. and Gilbert, W. in Grossman, L. and Moldave (eds.) (1980) Methods in Enzymology 65:499-560 Academic Press, New York.

An isolated nucleic acid encoding a human SOCS16 protein was identified. The nucleotide sequence and corresponding open reading frame are provided in SEQ ID NO: 1 through 4.

These SOCS16 proteins exhibit limited similarity to portions other cyclin associated proteins or transcription factors. In particular, β-sheet and α-

helix residues can be determined using, e.g., RASMOL program, see Sayle and Milner-White (1995) TIBS 20:374-376; or Gronenberg, et al. (1991) Protein Engineering 4:263-269; and other structural features are defined in Lodi, et al. (1994) Science 263:1762-1767.

This invention provides isolated DNA or fragments to encode a SOCS16 protein. In addition, this invention provides isolated or recombinant DNA which encodes a protein or polypeptide which is capable of hybridizing under appropriate conditions, e.g., high stringency, with the DNA sequences described herein. Said biologically active protein or polypeptide can be an intact protein, or fragment, and have an amino acid sequence as disclosed in SEQ ID NO: 2 or 4, particularly natural embodiments. Preferred embodiments will be full length natural sequences. Further, this invention contemplates the use of isolated or recombinant DNA, or fragments thereof, which encode proteins which are homologous to a SOCS16 protein or which were isolated using cDNA encoding a SOCS16 protein as a probe. The isolated DNA can have the respective regulatory sequences in the 5' and 3' flanks, e.g., promoters, enhancers, poly-A addition signals, and others. Also embraced are methods for making expression vectors with these sequences, or for making, e.g., expressing and purifying, protein products.

A DNA which codes for a SOCS16 protein will be particularly useful to identify genes, mRNA, and cDNA species which code for related or similar proteins, as well as DNAs which code for homologous proteins from different species. There are likely homologs in other species, including primates, rodents, canines, felines, and birds. Various SOCS16 proteins should be homologous and are encompassed herein. However, even proteins that have a more distant evolutionary relationship to the antigen can readily be isolated under appropriate conditions using these sequences if they are sufficiently

homologous. Primate SOCS16 proteins are of particular interest.

Recombinant clones derived from the genomic sequences, e.g., containing introns, will be useful for transgenic studies, including, e.g., transgenic cells and organisms, and for gene therapy. See, e.g., Goodnow (1992) "Transgenic Animals" in Roitt (ed.) Encyclopedia of Immunology, Academic Press, San Diego, pp. 1502-1504; Travis (1992) Science 256:1392-1394; Kuhn, et al. (1991) Science 254:707-710; Capecchi (1989) Science 244:1288; Robertson (1987)(ed.) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, IRL Press, Oxford; and Rosenberg (1992) J. Clinical Oncology 10:180-199.

IV. Antibodies

Antibodies can be raised to various SOCS16 proteins, including individual, polymorphic, allelic, strain, or species variants, and fragments thereof, both in their naturally occurring (full-length) forms and in their recombinant forms. Additionally, antibodies can be raised to SOCS16 proteins in either their active forms or in their inactive forms. Anti-idiotypic antibodies may also be used.

A. Antibody Production

A number of immunogens may be used to produce antibodies specifically reactive with SOCS16 proteins. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Naturally occurring protein may also be used either in pure or impure form. Synthetic peptides, made using the human SOCS16 protein sequences described herein, may also used as an immunogen for the production of antibodies to SOCS16 proteins. Recombinant protein can be expressed in eukaryotic or prokaryotic cells as described herein, and purified as described. Naturally folded or denatured material can be used, as appropriate, for producing antibodies. Either monoclonal or polyclonal antibodies

may be generated for subsequent use in immunoassays to measure the protein.

Methods of producing polyclonal antibodies are known to those of skill in the art. Typically, an immunogen, preferably a purified protein, is mixed with an adjuvant and animals are immunized with the mixture. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the SOCS16 protein of interest. When appropriately high titers of antibody to the immunogen are obtained, usually after repeated immunizations, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired. See, e.g., Harlow and Lane; or Coligan.

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Typically, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (see, Kohler and Milstein (1976) Eur. J. Immunol. 6:511-519, incorporated herein by reference). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according, e.g., to the general protocol outlined by Huse, et al. (1989) Science 246:1275-1281.

Antibodies, including binding fragments and single chain versions, against predetermined fragments of SOCS16

protein can be raised by immunization of animals with conjugates of the fragments with carrier proteins as described above. Monoclonal antibodies are prepared from cells secreting the desired antibody. These antibodies can be screened for binding to normal or defective SOCS16 proteins, or screened for agonistic or antagonistic activity, e.g., effect on cell cycle progression or transcription of specific genes. These monoclonal antibodies will usually bind with at least a K_D of about 1 mM, more usually at least about 300 µM, typically at least about 10 µM, more typically at least about 30 µM, preferably at least about 10 µM, and more preferably at least about 3 µM or better.

In some instances, it is desirable to prepare monoclonal antibodies from various mammalian hosts, such as mice, rodents, primates, humans, etc. Description of techniques for preparing such monoclonal antibodies may be found in, e.g., Stites, et al. (eds.) Basic and Clinical Immunology (4th ed.) Lange Medical Publications, Los Altos, CA, and references cited therein; Harlow and Lane (1988) Antibodies: A Laboratory Manual CSH Press; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.) Academic Press, New York, NY; and particularly in Kohler and Milstein (1975) Nature 256:495-497, which discusses one method of generating monoclonal antibodies. Summarized briefly, this method involves injecting an animal with an immunogen. The animal is then sacrificed and cells taken from its spleen, which are then fused with myeloma cells. The result is a hybrid cell or "hybridoma" that is capable of reproducing in vitro. The population of hybridomas is then screened to isolate individual clones, each of which secrete a single antibody species to the immunogen. In this manner, the individual antibody species obtained are the products of immortalized and cloned single B cells from the immune animal generated in response to a specific site recognized on the immunogenic substance.

Other suitable techniques involve selection of libraries of antibodies in phage or similar vectors. See, e.g., Huse, et al. (1989) "Generation of a Large Combinatorial Library of the Immunoglobulin Repertoire in Phage Lambda," Science 246:1275-1281; and Ward, et al. (1989) Nature 341:544-546. The polypeptides and antibodies of the present invention may be used with or without modification, including chimeric or humanized antibodies. Frequently, the polypeptides and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, chemiluminescent moieties, magnetic particles, and the like. Patents, teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced. See, Cabilly, U.S. Patent No. 4,816,567; and Queen, et al. (1989) Proc. Nat'l Acad. Sci. USA 86:10029-10033.

The antibodies of this invention are useful for affinity chromatography in isolating SOCS16 protein. Columns can be prepared where the antibodies are linked to a solid support, e.g., particles, such as agarose, SEPHADEX, or the like, where a cell lysate or supernatant may be passed through the column, the column washed, followed by increasing concentrations of a mild denaturant, whereby purified SOCS16 protein will be released.

The antibodies may also be used to screen expression libraries for particular expression products. Usually the antibodies used in such a procedure will be labeled with a moiety allowing easy detection of presence of antigen by antibody binding.

Antibodies to SOCS16 proteins may be used for the identification of cell populations expressing SOCS16 proteins. By assaying, e.g., by histology or otherwise, probably a disruptive assay which kills that sample of cells, the expression products of cells expressing SOCS16 proteins it is possible to diagnose disease, e.g., cancerous conditions.

Antibodies raised against each SOCS16 protein will also be useful to raise anti-idiotypic antibodies. These will be useful in detecting or diagnosing various immunological conditions related to expression of the respective antigens.

B. Immunoassays

A particular protein can be measured by a variety of immunoassay methods. For a review of immunological and immunoassay procedures in general, see Stites and Terr (eds.) (1991) Basic and Clinical Immunology (7th ed.). Moreover, the immunoassays of the present invention can be performed in many configurations, which are reviewed extensively in Maggio (ed.) (1980) Enzyme Immunoassay CRC Press, Boca Raton, Florida; Tijan (1985) "Practice and Theory of Enzyme Immunoassays," Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers B.V., Amsterdam; and Harlow and Lane Antibodies, A Laboratory Manual, supra, each of which is incorporated herein by reference. See also Chan (ed.) (1987) Immunoassay: A Practical Guide Academic Press, Orlando, FL; Price and Newman (eds.) (1991) Principles and Practice of Immunoassays Stockton Press, NY; and Ngo (ed.) (1988) Non-isotopic Immunoassays Plenum Press, NY.

Immunoassays for measurement of SOCS16 proteins can be performed by a variety of methods known to those skilled in the art. In brief, immunoassays to measure the protein can be either competitive or noncompetitive binding assays. In competitive binding assays, the sample to be analyzed competes with a labeled analyte for specific binding sites on a capture agent bound to a

solid surface. Preferably the capture agent is an antibody specifically reactive with SOCS16 proteins produced as described above. The concentration of labeled analyte bound to the capture agent is inversely proportional to the amount of free analyte present in the sample.

In a competitive binding immunoassay, the SOCS16 protein present in the sample competes with labeled protein for binding to a specific binding agent, for example, an antibody specifically reactive with the SOCS16 protein. The binding agent may be bound to a solid surface to effect separation of bound labeled protein from the unbound labeled protein. Alternately, the competitive binding assay may be conducted in liquid phase and a variety of techniques known in the art may be used to separate the bound labeled protein from the unbound labeled protein. Following separation, the amount of bound labeled protein is determined. The amount of protein present in the sample is inversely proportional to the amount of labeled protein binding.

Alternatively, a homogeneous immunoassay may be performed in which a separation step is not needed. In these immunoassays, the label on the protein is altered by the binding of the protein to its specific binding agent. This alteration in the labeled protein results in a decrease or increase in the signal emitted by label, so that measurement of the label at the end of the immunoassay allows for detection or quantitation of the protein.

SOCS16 proteins may also be determined by a variety of noncompetitive immunoassay methods. For example, a two-site, solid phase sandwich immunoassay may be used. In this type of assay, a binding agent for the protein, for example an antibody, is attached to a solid support. A second protein binding agent, which may also be an antibody, and which binds the protein at a different site, is labeled. After binding at both sites on the

protein has occurred, the unbound labeled binding agent is removed and the amount of labeled binding agent bound to the solid phase is measured. The amount of labeled binding agent bound is directly proportional to the amount of protein in the sample.

Western blot analysis can be used to determine the presence of SOCS16 proteins in a sample. Electrophoresis is carried out, for example, on a tissue sample suspected of containing the protein. Following electrophoresis to separate the proteins, and transfer of the proteins to a suitable solid support, e.g., a nitrocellulose filter, the solid support is incubated with an antibody reactive with the protein. This antibody may be labeled, or alternatively may be detected by subsequent incubation with a second labeled antibody that binds the primary antibody.

The immunoassay formats described above employ labeled assay components. The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. A wide variety of labels and methods may be used. Traditionally, a radioactive label incorporating ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P was used. Non-radioactive labels include proteins which bind to labeled antibodies, fluorophores, chemiluminescent agents, enzymes, and antibodies which can serve as specific binding pair members for a labeled protein. The choice of label depends on sensitivity required, ease of conjugation with the compound, stability requirements, and available instrumentation. For a review of various labeling or signal producing systems which may be used, see U.S. Patent No. 4,391,904, which is incorporated herein by reference.

Antibodies reactive with a particular protein can also be measured by a variety of immunoassay methods. For a review of immunological and immunoassay procedures applicable to the measurement of antibodies by

immunoassay techniques, see Stites and Terr (eds.) Basic and Clinical Immunology (7th ed.) supra; Maggio (ed.) Enzyme Immunoassay, supra; and Harlow and Lane Antibodies, A Laboratory Manual, supra.

In brief, immunoassays to measure antisera reactive with SOCS16 proteins can be either competitive or noncompetitive binding assays. In competitive binding assays, the sample analyte competes with a labeled analyte for specific binding sites on a capture agent bound to a solid surface. Preferably the capture agent is a purified recombinant SOCS16 protein produced as described above. Other sources of SOCS16 proteins, including isolated or partially purified naturally occurring protein, may also be used. Noncompetitive assays include sandwich assays, in which the sample analyte is bound between two analyte-specific binding reagents. One of the binding agents is used as a capture agent and is bound to a solid surface. The second binding agent is labeled and is used to measure or detect the resultant complex by visual or instrument means. A number of combinations of capture agent and labeled binding agent can be used. A variety of different immunoassay formats, separation techniques, and labels can be also be used similar to those described above for the measurement of SOCS16 proteins.

V. Making SOCS16 proteins; Mimetics

DNAs which encode a SOCS16 protein or fragments thereof can be obtained by chemical synthesis, screening cDNA libraries, or by screening genomic libraries prepared from a wide variety of cell lines or tissue samples. Methods for doing so, or making expression vectors are described herein.

These DNAs can be expressed in a wide variety of host cells for the synthesis of a full-length protein or fragments which can in turn, e.g., be used to generate polyclonal or monoclonal antibodies; for binding studies;

for construction and expression of modified molecules; and for structure/function studies. Each SOCS16 protein or its fragments can be expressed in host cells that are transformed or transfected with appropriate expression vectors. These molecules can be substantially purified to be free of protein or cellular contaminants, other than those derived from the recombinant host, and therefore are particularly useful in pharmaceutical compositions when combined with a pharmaceutically acceptable carrier and/or diluent. The antigen, e.g., SOCS16 protein, or portions thereof, may be expressed as fusions with other proteins or possessing an epitope tag.

Expression vectors are typically self-replicating DNA or RNA constructs containing the desired antigen gene or its fragments, usually operably linked to appropriate genetic control elements that are recognized in a suitable host cell. The specific type of control elements necessary to effect expression will depend upon the eventual host cell used. Generally, the genetic control elements can include a prokaryotic promoter system or a eukaryotic promoter expression control system, and typically include a transcriptional promoter, an optional operator to control the onset of transcription, transcription enhancers to elevate the level of mRNA expression, a sequence that encodes a suitable ribosome binding site, and sequences that terminate transcription and translation. Expression vectors also usually contain an origin of replication that allows the vector to replicate independently from the host cell.

The vectors of this invention contain DNAs which encode a SOCS16 protein, or a fragment thereof, typically encoding, e.g., a biologically active polypeptide, or protein. The DNA can be under the control of a viral promoter and can encode a selection marker. This invention further contemplates use of such expression vectors which are capable of expressing eukaryotic cDNA

coding for a SOCS16 protein in a prokaryotic or eukaryotic host, where the vector is compatible with the host and where the eukaryotic cDNA coding for the protein is inserted into the vector such that growth of the host containing the vector expresses the cDNA in question. Usually, expression vectors are designed for stable replication in their host cells or for amplification to greatly increase the total number of copies of the desirable gene per cell. It is not always necessary to require that an expression vector replicate in a host cell, e.g., it is possible to effect transient expression of the protein or its fragments in various hosts using vectors that do not contain a replication origin that is recognized by the host cell. It is also possible to use vectors that cause integration of a SOCS16 protein gene or its fragments into the host DNA by recombination, or to integrate a promoter which controls expression of an endogenous gene.

Vectors, as used herein, contemplate plasmids, viruses, bacteriophage, integratable DNA fragments, and other vehicles which enable the integration of DNA fragments into the genome of the host. Expression vectors are specialized vectors which contain genetic control elements that effect expression of operably linked genes. Plasmids are the most commonly used form of vector, but many other forms of vectors which serve an equivalent function are suitable for use herein. See, e.g., Pouwels, et al. (1985 and Supplements) Cloning Vectors: A Laboratory Manual Elsevier, N.Y.; and Rodriquez, et al. (eds.) (1988) Vectors: A Survey of Molecular Cloning Vectors and Their Uses Buttersworth, Boston, MA.

Suitable host cells include prokaryotes, lower eukaryotes, and higher eukaryotes. Prokaryotes include both gram negative and gram positive organisms, e.g., E. coli and B. subtilis. Lower eukaryotes include yeasts, e.g., S. cerevisiae and Pichia, and species of the genus

Dictyostelium. Higher eukaryotes include established tissue culture cell lines from animal cells, both of non-mammalian origin, e.g., insect cells, and birds, and of mammalian origin, e.g., human, primates, and rodents.

Prokaryotic host-vector systems include a wide variety of vectors for many different species. As used herein, E. coli and its vectors will be used generically to include equivalent vectors used in other prokaryotes. A representative vector for amplifying DNA is pBR322 or its derivatives. Vectors that can be used to express SOCS16 proteins or SOCS16 protein fragments include, but are not limited to, such vectors as those containing the lac promoter (pUC-series); trp promoter (pBR322-trp); Ipp promoter (the pIN-series); lambda-pP or pR promoters (pOTS); or hybrid promoters such as ptac (pDR540). See Brosius, et al. (1988) "Expression Vectors Employing Lambda-, trp-, lac-, and Ipp-derived Promoters", in Rodriguez and Denhardt (eds.) Vectors: A Survey of Molecular Cloning Vectors and Their Uses 10:205-236 Buttersworth, Boston, MA.

Lower eukaryotes, e.g., yeasts and Dictyostelium, may be transformed with SOCS16 protein sequence containing vectors. For purposes of this invention, the most common lower eukaryotic host is the baker's yeast, Saccharomyces cerevisiae. It will be used generically to represent lower eukaryotes although a number of other strains and species are also available. Yeast vectors typically consist of a replication origin (unless of the integrating type), a selection gene, a promoter, DNA encoding the desired protein or its fragments, and sequences for translation termination, polyadenylation, and transcription termination. Suitable expression vectors for yeast include such constitutive promoters as 3-phosphoglycerate kinase and various other glycolytic enzyme gene promoters or such inducible promoters as the alcohol dehydrogenase 2 promoter or metallothionine promoter. Suitable vectors include derivatives of the

following types: self-replicating low copy number (such as the YRp-series), self-replicating high copy number (such as the YEp-series); integrating types (such as the YIp-series), or mini-chromosomes (such as the YCp-series).

Higher eukaryotic tissue culture cells are typically the preferred host cells for expression of the functionally active SOCS16 protein. In principle, many higher eukaryotic tissue culture cell lines may be used, e.g., insect baculovirus expression systems, whether from an invertebrate or vertebrate source. However, mammalian cells are preferred to achieve proper processing, both cotranslationally and posttranslationally. Transformation or transfection and propagation of such cells is routine. Useful cell lines include HeLa cells, Chinese hamster ovary (CHO) cell lines, baby rat kidney (BRK) cell lines, insect cell lines, bird cell lines, and monkey (COS) cell lines. Expression vectors for such cell lines usually include an origin of replication, a promoter, a translation initiation site, RNA splice sites (e.g., if genomic DNA is used), a polyadenylation site, and a transcription termination site. These vectors also may contain a selection gene or amplification gene. Suitable expression vectors may be plasmids, viruses, or retroviruses carrying promoters derived, e.g., from such sources as from adenovirus, SV40, parvoviruses, vaccinia virus, or cytomegalovirus. Representative examples of suitable expression vectors include pCDNA1; pCD, see Okayama, et al. (1985) Mol. Cell Biol. 5:1136-1142; pMC1neo Poly-A, see Thomas, et al. (1987) Cell 51:503-512; and a baculovirus vector such as pAC 373 or pAC 610.

It is likely that SOCS16 proteins need not be glycosylated to elicit biological responses. However, it will occasionally be desirable to express a SOCS16 protein polypeptide in a system which provides a specific or defined glycosylation pattern. In this case, the usual pattern will be that provided naturally by the

expression system. However, the pattern will be modifiable by exposing the polypeptide, e.g., in unglycosylated form, to appropriate glycosylating proteins introduced into a heterologous expression system. For example, the SOCS16 protein gene may be co-transformed with one or more genes encoding mammalian or other glycosylating enzymes. It is further understood that over glycosylation may be detrimental to SOCS16 protein biological activity, and that one of skill may perform routine testing to optimize the degree of glycosylation which confers optimal biological activity.

A SOCS16 protein, or a fragment thereof, may be engineered to be phosphatidyl inositol (PI) linked to a cell membrane, but can be removed from membranes by treatment with a phosphatidyl inositol cleaving enzyme, e.g., phosphatidyl inositol phospholipase-C. This releases the antigen in a biologically active form, and allows purification by standard procedures of protein chemistry. See, e.g., Low (1989) Biochem. Biophys. Acta 988:427-454; Tse, et al. (1985) Science 230:1003-1008; and Brunner, et al. (1991) J. Cell Biol. 114:1275-1283.

Now that SOCS16 proteins have been characterized, fragments or derivatives thereof can be prepared by conventional processes for synthesizing peptides. These include processes such as are described in Stewart and Young (1984) Solid Phase Peptide Synthesis Pierce Chemical Co., Rockford, IL; Bodanszky and Bodanszky (1984) The Practice of Peptide Synthesis Springer-Verlag, New York, NY; and Bodanszky (1984) The Principles of Peptide Synthesis Springer-Verlag, New York, NY. For example, an azide process, an acid chloride process, an acid anhydride process, a mixed anhydride process, an active ester process (for example, p-nitrophenyl ester, N-hydroxysuccinimide ester, or cyanomethyl ester), a carbodiimidazole process, an oxidative-reductive process, or a dicyclohexylcarbodiimide (DCCD)/additive process can

be used. Solid phase and solution phase syntheses are both applicable to the foregoing processes.

The prepared protein and fragments thereof can be isolated and purified from the reaction mixture by means of peptide separation, for example, by extraction, precipitation, electrophoresis and various forms of chromatography, and the like. The SOCS16 proteins of this invention can be obtained in varying degrees of purity depending upon its desired use. Purification can be accomplished by use of known protein purification techniques or by the use of the antibodies or binding partners herein described, e.g., in immunoabsorbant affinity chromatography. This immunoabsorbant affinity chromatography is carried out by first linking the antibodies to a solid support and then contacting the linked antibodies with solubilized lysates of appropriate source cells, lysates of other cells expressing the protein, or lysates or supernatants of cells producing the SOCS16 proteins as a result of recombinant DNA techniques, see below.

Multiple cell lines may be screened for one which expresses a SOCS16 protein at a high level compared with other cells. Various cell lines, e.g., a mouse thymic stromal cell line TA4, is screened and selected for its favorable handling properties. Natural SOCS16 proteins can be isolated from natural sources, or by expression from a transformed cell using an appropriate expression vector. Purification of the expressed protein is achieved by standard procedures, or may be combined with engineered means for effective purification at high efficiency from cell lysates or supernatants. Epitope or other tags, e.g., FLAG or His_6 segments, can be used for such purification features.

VI. Physical Variants

This invention also encompasses proteins or peptides having substantial amino acid sequence similarity with an

amino acid sequence of a SOCS16 protein. Natural variants include individual, polymorphic, allelic, strain, or species variants.

Amino acid sequence similarity, or sequence identity, is determined by optimizing residue matches, if necessary, by introducing gaps as required. This changes when considering conservative substitutions as matches. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. Homologous amino acid sequences include natural polymorphic, allelic, and interspecies variations in each respective protein sequence. Typical homologous proteins or peptides will have from 50-100% similarity (if gaps can be introduced), to 75-100% similarity (if conservative substitutions are included) over fixed stretches of amino acids with the amino acid sequence of the SOCS16 protein. Similarity measures will be at least about 50%, generally at least 65%, usually at least 70%, preferably at least 75%, and more preferably at least 90%, and in particularly preferred embodiments, at least 96% or more. See also Needleham, et al. (1970) J. Mol. Biol. 48:443-453; Sankoff, et al. (1983) Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison Chapter One, Addison-Wesley, Reading, MA; and software packages from IntelliGenetics, Mountain View, CA; and the University of Wisconsin Genetics Computer Group, Madison, WI. Stretches of amino acids will be at least about 10 amino acids, usually about 20 amino acids, usually 50 amino acids, preferably 75 amino acids, and in particularly preferred embodiments at least about 100 amino acids. Identity can also be measures over amino acid stretches of about 98, 99, 110, 120, 130, etc.

Nucleic acids encoding mammalian SOCS16 proteins will typically hybridize to the nucleic acid sequence of the coding portions of SEQ ID NO: 1 or 3 under stringent conditions. For example, nucleic acids encoding human SOCS16 proteins will normally hybridize to the nucleic acid of SEQ ID NO: 1 or 3 under stringent hybridization conditions. Generally, stringent conditions are selected to be about 10° C lower than the thermal melting point (Tm) for the probe sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent conditions will be those in which the salt concentration is about 0.2 molar at pH 7 and the temperature is at least about 50° C. Other factors may significantly affect the stringency of hybridization, including, among others, base composition and size of the complementary strands, the presence of organic solvents such as formamide, and the extent of base mismatching. A preferred embodiment will include nucleic acids which will bind to disclosed sequences in 50% formamide and 200 mM NaCl at 42° C.

An isolated SOCS16 protein DNA can be readily modified by nucleotide substitutions, nucleotide deletions, nucleotide insertions, and short inversions of nucleotide stretches. These modifications result in novel DNA sequences which encode SOCS16 protein antigens, their derivatives, or proteins having highly similar physiological, immunogenic, or antigenic activity.

Modified sequences can be used to produce mutant antigens or to enhance expression. Enhanced expression may involve gene amplification, increased transcription, increased translation, and other mechanisms. Such mutant SOCS16 protein derivatives include predetermined or site-specific mutations of the respective protein or its fragments. "Mutant SOCS16 protein" encompasses a polypeptide otherwise falling within the homology

definition of the human SOCS16 protein as set forth above, but having an amino acid sequence which differs from that of a SOCS16 protein as found in nature, whether by way of deletion, substitution, or insertion. In particular, "site specific mutant SOCS16 protein" generally includes proteins having significant similarity with a protein having a sequence of SEQ ID NO: 2 or 4, e.g., natural embodiments, and as sharing various biological activities, e.g., antigenic or immunogenic, with those sequences, and in preferred embodiments contain most or all of the disclosed sequence. This applies also to polymorphic variants from different individuals. Similar concepts apply to different SOCS16 proteins, particularly those found in various warm blooded animals, e.g., mammals and birds. As stated before, it is emphasized that descriptions are generally meant to encompass other SOCS16 proteins, not limited to the human embodiments specifically discussed.

Although site specific mutation sites are predetermined, mutants need not be site specific. SOCS16 protein mutagenesis can be conducted by making amino acid insertions or deletions. Substitutions, deletions, insertions, or any combinations may be generated to arrive at a final construct. Insertions include amino- or carboxyl- terminal fusions, e.g. epitope tags. Random mutagenesis can be conducted at a target codon and the expressed mutants can then be screened for the desired activity. Methods for making substitution mutations at predetermined sites in DNA having a known sequence are well known in the art, e.g., by M13 primer mutagenesis or polymerase chain reaction (PCR) techniques. See also, Sambrook, et al. (1989) and Ausubel, et al. (1987 and Supplements). The mutations in the DNA normally should not place coding sequences out of reading frames and preferably will not create complementary regions that could hybridize to produce secondary mRNA structure such as loops or hairpins.

The present invention also provides recombinant proteins, e.g., heterologous fusion proteins using segments from these proteins. A heterologous fusion protein is a fusion of proteins or segments which are naturally not normally fused in the same manner. Thus, the fusion product of an immunoglobulin with a SOCS16 protein polypeptide is a continuous protein molecule having sequences fused in a typical peptide linkage, typically made as a single translation product and exhibiting properties derived from each source peptide. A similar concept applies to heterologous nucleic acid sequences.

In addition, new constructs may be made from combining similar functional domains from other proteins. For example, protein-binding or other segments may be "swapped" between different new fusion polypeptides or fragments. See, e.g., Cunningham, et al. (1989) Science 243:1330-1336; and O'Dowd, et al. (1988) J. Biol. Chem. 263:15985-15992. Thus, new chimeric polypeptides exhibiting new combinations of specificities will result from the functional linkage of protein-binding specificities and other functional domains.

VII. Functional Variants

The blocking of physiological response to SOCS16 protein may result from the inhibition of binding of the protein to its binding partner, e.g., through competitive inhibition. Thus, in vitro assays of the present invention will often use isolated protein, membranes from cells expressing a recombinant membrane associated SOCS16 protein, soluble fragments comprising binding segments of these proteins, or fragments attached to solid phase substrates. These assays will also allow for the diagnostic determination of the effects of either binding segment mutations and modifications, or protein mutations and modifications, e.g., protein analogs. This invention also contemplates the use of competitive drug screening

assays, e.g., where neutralizing antibodies to antigen or binding partner fragments compete with a test compound for binding to the protein. In this manner, the antibodies can be used to detect the presence of a polypeptide which shares one or more antigenic binding sites of the protein and can also be used to occupy binding sites on the protein that might otherwise interact with a binding partner.

"Derivatives" of SOCS16 protein antigens include amino acid sequence mutants, glycosylation variants, and covalent or aggregate conjugates with other chemical moieties. Covalent derivatives can be prepared by linkage of functionalities to groups which are found in SOCS16 protein amino acid side chains or at the N- or C-termini, by means which are well known in the art. These derivatives can include, without limitation, aliphatic esters or amides of the carboxyl terminus, or of residues containing carboxyl side chains, O-acyl derivatives of hydroxyl group-containing residues, and N-acyl derivatives of the amino terminal amino acid or amino-group containing residues, e.g., lysine or arginine. Acyl groups are selected from the group of alkyl-moieties including C3 to C18 normal alkyl, thereby forming alkanoyl aroyl species. Covalent attachment to carrier proteins may be important when immunogenic moieties are haptens.

In particular, glycosylation alterations are included, e.g., made by modifying the glycosylation patterns of a polypeptide during its synthesis and processing, or in further processing steps. Particularly preferred means for accomplishing this are by exposing the polypeptide to glycosylating enzymes derived from cells which normally provide such processing, e.g., mammalian glycosylation enzymes. Deglycosylation enzymes are also contemplated. Also embraced are versions of the same primary amino acid sequence which have other minor modifications, including phosphorylated amino acid

residues, e.g., phosphotyrosine, phosphoserine, or phosphothreonine, or other moieties, including ribosyl groups or cross-linking reagents.

A major group of derivatives are covalent conjugates of the SOCS16 protein or fragments thereof with other proteins or polypeptides. These derivatives can be synthesized in recombinant culture such as N- or C-terminal fusions or by the use of agents known in the art for their usefulness in cross-linking proteins through reactive side groups. Preferred protein derivatization sites with cross-linking agents are at free amino groups, carbohydrate moieties, and cysteine residues.

Fusion polypeptides between SOCS16 protein and other homologous or heterologous proteins are also provided. Heterologous polypeptides may be fusions between different surface markers, resulting in, e.g., a hybrid protein exhibiting binding partner specificity. Likewise, heterologous fusions may be constructed which would exhibit a combination of properties or activities of the derivative proteins. Typical examples are fusions of a reporter polypeptide, e.g., luciferase, with a segment or domain of a protein, e.g., a segment involved in binding partner interaction, so that the presence or location of the fused protein may be easily determined. See, e.g., Dull, et al., U.S. Patent No. 4,859,609. Other gene fusion partners include bacterial β-galactosidase, trpE, Protein A, ß-lactamase, alpha amylase, alcohol dehydrogenase, and yeast alpha mating factor. See, e.g., Godowski, et al. (1988) Science 241:812-816. The fusion partner can be constructed such that it can be cleaved off such that a protein of substantially natural length is generated.

Such polypeptides may also have amino acid residues which have been chemically modified by phosphorylation, sulfonation, biotinylation, or the addition or removal of other moieties, particularly those which have molecular shapes similar to phosphate groups. In some embodiments,

the modifications will be useful labeling reagents, or serve as purification targets, e.g., affinity proteins.

This invention also contemplates the use of derivatives of SOCS16 protein other than variations in amino acid sequence or glycosylation. Such derivatives may involve covalent or aggregative association with chemical moieties. These derivatives generally fall into the three classes: (1) salts, (2) side chain and terminal residue covalent modifications, and (3) adsorption complexes, for example with cell membranes. Such covalent or aggregative derivatives are useful as immunogens, as reagents in immunoassays, or in purification methods such as for affinity purification of proteinss or other binding proteins. For example, a SOCS16 protein antigen can be immobilized by covalent bonding to a solid support such as cyanogen bromide-activated SEPHAROSE, by methods which are well known in the art, or adsorbed onto polyolefin surfaces, with or without glutaraldehyde cross-linking, for use in the assay or purification of anti-SOCS16 protein antibodies or its respective binding partner. The SOCS16 protein can also be labeled with a detectable group, e.g., radioiodinated by the chloramine T procedure, covalently bound to rare earth chelates, or conjugated to another fluorescent moiety for use in diagnostic assays. Purification of SOCS16 proteins may be effected by immobilized antibodies or binding partner.

Isolated SOCS16 protein genes will allow transformation of cells lacking expression of corresponding SOCS16 protein, e.g., either species types or cells which lack corresponding proteins and exhibit negative background activity. Expression of transformed genes will allow isolation of antigenically pure cell lines, with defined or single specie variants. This approach will allow for more sensitive detection and discrimination of the physiological effects of SOCS16 binding proteins. Subcellular fragments, e.g.,

cytoplasts or membrane fragments, can be isolated and used.

VIII. Binding Agent:SOCS16 Protein Complexes

A SOCS16 protein that specifically binds to or that is specifically immunoreactive with an antibody generated against a defined immunogen, such as an immunogen consisting of the amino acid sequence of SEQ ID NO: 2 or 4 is typically determined in an immunoassay. The immunoassay uses a polyclonal antiserum which was raised to a protein of SEQ ID NO: 2 or 4. This antiserum is selected to have low crossreactivity against other intracellular regulatory proteins of the SOCS family and any such crossreactivity is removed by immunoabsorbtion prior to use in the immunoassay.

In order to produce antisera for use in an immunoassay, the protein of SEQ ID NO: 2 or 4 is isolated as described herein. For example, recombinant protein may be produced in a mammalian cell line. An inbred strain of mice such as/c is immunized with the protein of SEQ ID NO: 2 or 4 using a standard adjuvant, such as Freund's adjuvant, and a standard mouse immunization protocol (see Harlow and Lane, supra). Alternatively, a synthetic peptide, preferably near full length, derived from the sequences disclosed herein and conjugated to a carrier protein can be used an immunogen. Polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against other intracellular proteins, using a competitive binding immunoassay such as the one described in Harlow and Lane, supra, at pages 570-573. Preferably two intracellular proteins are used in this determination in conjunction with human SOCS16 protein.

Immunoassays in the competitive binding format can be used for the crossreactivity determinations. For example, a protein of SEQ ID NO: 2 can be immobilized to a solid support. Proteins added to the assay compete with the binding of the antisera to the immobilized antigen. The ability of the above proteins to compete with the binding of the antisera to the immobilized protein is compared to the protein of SEQ ID NO: 2. The percent crossreactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% crossreactivity with each of the proteins listed above are selected and pooled. The cross-reacting antibodies are then removed from the pooled antisera by immunoabsorbtion with the above-listed proteins.

The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein to the immunogen protein (e.g., the SOCS16 protein of SEQ ID NO: 2). In order to make this comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. If the amount of the second protein required is less than twice the amount of the protein, e.g., of SEQ ID NO: 2 that is required, then the second protein is said to specifically bind to an antibody generated to the immunogen.

It is understood that each of SOCS16 proteins are members of respective families of homologous proteins that comprise two or more genes. For a particular gene product, such as the human SOCS16 protein, the term refers not only to the amino acid sequences disclosed herein, but also to other proteins that are polymorphic, allelic, non-allelic, or species variants. It is also understood that the term "SOCS16 protein" includes nonnatural mutations introduced by deliberate mutation

using conventional recombinant technology such as single site mutation, or by excising short sections of DNA encoding SOCS16 proteins, or by substituting new amino acids, or adding new amino acids. Such minor alterations should substantially maintain the immunoidentity of the original molecule and/or its biological activity. Thus, these alterations include proteins that are specifically immunoreactive with a designated naturally occurring SOCS16 protein, for example, the human SOCS16 protein shown in SEQ ID NO: 4. The biological properties of the altered proteins can be determined by expressing the protein in an appropriate cell line and measuring, e.g., a proliferative effect. Particular protein modifications considered minor would include conservative substitution of amino acids with similar chemical properties, as described above for the SOCS16 protein as a whole. By aligning a protein optimally with the protein of SEQ ID NO: 2 or 4, and by using the conventional immunoassays described herein to determine immunoidentity, or by using proliferative assays, one can determine the protein compositions of the invention.

X. Uses

The present invention provides reagents which will find use in diagnostic applications as described elsewhere herein, e.g., in the general description for developmental abnormalities, or below in the description of kits for diagnosis.

SOCS16 nucleotides, e.g., human SOCS16 DNA or RNA, may be used as a component in a forensic assay. For instance, the nucleotide sequences provided may be labeled using, e.g., ^{32}P or biotin and used to probe standard restriction fragment polymorphism blots, providing a measurable character to aid in distinguishing between individuals. Such probes may be used in well-known forensic techniques such as genetic fingerprinting. In addition, nucleotide probes made from SOCS16 sequences

may be used in in situ assays to detect chromosomal abnormalities. For instance, rearrangements in the human chromosome encoding a SOCS16 gene may be detected via well-known in situ techniques, using SOCS16 probes in conjunction with other known chromosome markers.

Antibodies and other binding agents directed towards SOCS16 proteins or nucleic acids may be used to purify the corresponding SOCS16 molecule. As described in the Examples below, antibody purification of SOCS16 protein components is both possible and practicable. Antibodies and other binding agents may also be used in a diagnostic fashion to determine whether SOCS16 protein components are present in a tissue sample or cell population using well-known techniques described herein. The ability to attach a binding agent to a SOCS16 protein provides a means to diagnose disorders associated with SOCS16 protein misregulation. Antibodies and other SOCS16 protein binding agents may also be useful as histological markers. As described in the examples below, SOCS16 protein expression is limited to specific tissue types. By directing a probe, such as an antibody or nucleic acid to a SOCS16 protein it is possible to use the probe to distinguish tissue and cell types in situ or in vitro.

This invention also provides reagents with significant therapeutic value. The SOCS16 protein (naturally occurring or recombinant), fragments thereof, and antibodies thereto, along with compounds identified as having binding affinity to a SOCS16 protein, are useful in the treatment of conditions associated with abnormal physiology or development, including abnormal proliferation, e.g., cancerous conditions, or degenerative conditions. Abnormal proliferation, regeneration, degeneration, and atrophy may be modulated by appropriate therapeutic treatment using the compositions provided herein. For example, a disease or disorder associated with abnormal expression or abnormal signaling by a SOCS16 protein is a target for an agonist

or antagonist of the protein. The proteins likely play a role in regulation or development of neuronal or hematopoietic cells, e.g., lymphoid cells, which affect immunological responses.

Other abnormal developmental conditions are known in cell types shown to possess SOCS16 protein mRNA by northern blot analysis. See Berkow (ed.) The Merck Manual of Diagnosis and Therapy, Merck & Co., Rahway, NJ; and Thorn, et al. Harrison's Principles of Internal Medicine, McGraw-Hill, NY. Developmental or functional abnormalities, e.g., of the neuronal or immune system, cause significant medical abnormalities and conditions which may be susceptible to prevention or treatment using compositions provided herein.

Recombinant SOCS16 protein or SOCS16 protein antibodies can be purified and then administered to a patient. These reagents can be combined for therapeutic use with additional active or inert ingredients, e.g., in conventional pharmaceutically acceptable carriers or diluents, e.g., immunogenic adjuvants, along with physiologically innocuous stabilizers and excipients. These combinations can be sterile filtered and placed into dosage forms as by lyophilization in dosage vials or storage in stabilized aqueous preparations. This invention also contemplates use of antibodies or binding fragments thereof, including forms which are not complement binding.

Drug screening using antibodies or fragments thereof can identify compounds having binding affinity to SOCS16 protein, including isolation of associated components. Subsequent biological assays can then be utilized to determine if the compound has intrinsic stimulating activity and is therefore a blocker or antagonist in that it blocks the activity of the protein. Likewise, a compound having intrinsic stimulating activity can activate the binding partner and is thus an agonist in that it simulates the activity of a SOCS16 protein. This

invention further contemplates the therapeutic use of antibodies to SOCS16 protein as antagonists. This approach should be particularly useful with other SOCS16 protein species variants.

The quantities of reagents necessary for effective therapy will depend upon many different factors, including means of administration, target site, physiological state of the patient, and other medicants administered. Thus, treatment dosages should be titrated to optimize safety and efficacy. Typically, dosages used in vitro may provide useful guidance in the amounts useful for in situ administration of these reagents. Animal testing of effective doses for treatment of particular disorders will provide further predictive indication of human dosage. Various considerations are described, e.g., in Gilman, et al. (eds.) (1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics (8th ed.) Pergamon Press; and (1990) Remington's Pharmaceutical Sciences (17th ed.) Mack Publishing Co., Easton, PA. Methods for administration are discussed therein and below, e.g., for oral, intravenous, intraperitoneal, or intramuscular administration, transdermal diffusion, and others. Pharmaceutically acceptable carriers will include water, saline, buffers, and other compounds described, e.g., in the Merck Index, Merck & Co., Rahway, NJ. Dosage ranges would ordinarily be expected to be in amounts lower than 1 mM concentrations, typically less than about 10 µM concentrations, usually less than about 100 nM, preferably less than about 10 pM (picomolar), and most preferably less than about 1 fM (femtomolar), with an appropriate carrier. Slow release formulations, or a slow release apparatus will often be utilized for continuous administration.

SOCS16 protein, fragments thereof, and antibodies to it or its fragments, antagonists, and agonists, may be administered directly to the host to be treated or,

depending on the size of the compounds, it may be desirable to conjugate them to carrier proteins such as ovalbumin or serum albumin prior to their administration. Therapeutic formulations may be administered in any conventional dosage formulation. While it is possible for the active ingredient to be administered alone, it is preferable to present it as a pharmaceutical formulation. Formulations typically comprise at least one active ingredient, as defined above, together with one or more acceptable carriers thereof. Each carrier should be both pharmaceutically and physiologically acceptable in the sense of being compatible with the other ingredients and not injurious to the patient. Formulations include those suitable for oral, rectal, nasal, or parenteral (including subcutaneous, intramuscular, intravenous and intradermal) administration. The formulations may conveniently be presented in unit dosage form and may be prepared by any methods well known in the art of pharmacy. See, e.g., Gilman, et al. (eds.) (1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics (8th ed.) Pergamon Press; and (1990) Remington's Pharmaceutical Sciences (17th ed.) Mack Publishing Co., Easton, PA; Avis, et al. (eds.) (1993) Pharmaceutical Dosage Forms: Parenteral Medications Dekker, NY; Lieberman, et al. (eds.) (1990) Pharmaceutical Dosgae Forms: Tablets Dekker, NY; and Lieberman, et al. (eds.) (1990) Pharmaceutical Dosage Forms: Disperse Systems Dekker, NY. The therapy of this invention may be combined with or used in association with other therapeutic agents.

Both the naturally occurring and the recombinant forms of the SOCS16 protein of this invention are particularly useful in kits and assay methods which are capable of screening compounds for binding activity to the proteins. Several methods of automating assays have been developed in recent years so as to permit screening of tens of thousands of compounds in a short period.

See, e.g., Fodor, et al. (1991) Science 251:767-773, and other descriptions of chemical diversity libraries, which describe means for testing of binding affinity by a plurality of compounds. The development of suitable assays can be greatly facilitated by the availability of large amounts of purified, soluble SOCS16 protein as provided by this invention.

For example, antagonists can normally be found once the protein has been structurally defined. Testing of potential protein analogs is now possible upon the development of highly automated assay methods using a purified binding partner. In particular, new agonists and antagonists will be discovered by using screening techniques described herein. Of particular importance are compounds found to have a combined binding affinity for multiple SOCS16 protein binding components, e.g., compounds which can serve as antagonists for species variants of a SOCS16 protein.

This invention is particularly useful for screening compounds by using recombinant protein in a variety of drug screening techniques. The advantages of using a recombinant protein in screening for specific binding partners include: (a) improved renewable source of the SOCS16 protein from a specific source; (b) potentially greater number of binding partners per cell giving better signal to noise ratio in assays; and (c) species variant specificity (theoretically giving greater biological and disease specificity).

One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant DNA molecules expressing a SOCS16 protein binding counterpart. Cells may be isolated which express a binding counterpart in isolation from any others. Such cells, either in viable or fixed form, can be used for standard protein binding assays. See also, Parce, et al. (1989) Science 246:243-247; and Owicki, et al. (1990) Proc. Nat'l Acad. Sci. USA 87:4007-4011, which describe

sensitive methods to detect cellular responses. Competitive assays are particularly useful, where the cells (source of SOCS16 protein) are contacted and incubated with a labeled binding partner or antibody having known binding affinity to the protein, such as ^{125}I-antibody, and a test sample whose binding affinity to the binding composition is being measured. The bound and free labeled binding compositions are then separated to assess the degree of protein binding. The amount of test compound bound is inversely proportional to the amount of labeled binding partner binding to the known source. Any one of numerous techniques can be used to separate bound from free protein to assess the degree of protein binding. This separation step could typically involve a procedure such as adhesion to filters followed by washing, adhesion to plastic followed by washing, or centrifugation of the cell membranes. Viable cells could also be used to screen for the effects of drugs on SOCS16 protein mediated functions, e.g., second messenger levels, i.e., cell proliferation; inositol phosphate pool changes, transcription using a luciferase-type assay; and others. Some detection methods allow for elimination of a separation step, e.g., a proximity sensitive detection system.

Another method utilizes membranes from transformed eukaryotic or prokaryotic host cells as the source of a SOCS16 protein. These cells are stably transformed with DNA vectors directing the expression of a SOCS16 protein, e.g., an engineered membrane bound form. Essentially, the membranes would be prepared from the cells and used in a protein binding assay such as the competitive assay set forth above.

Still another approach is to use solubilized, unpurified or solubilized, purified SOCS16 protein from transformed eukaryotic or prokaryotic host cells. This allows for a "molecular" binding assay with the

advantages of increased specificity, the ability to automate, and high drug test throughput.

Another technique for drug screening involves an approach which provides high throughput screening for compounds having suitable binding affinity to a SOCS16 protein antibody and is described in detail in Geysen, European Patent Application 84/03564, published on September 13, 1984. First, large numbers of different small peptide test compounds are synthesized on a solid substrate, e.g., plastic pins or some other appropriate surface, see Fodor, et al., supra. Then all the pins are reacted with solubilized, unpurified or solubilized, purified SOCS16 protein antibody, and washed. The next step involves detecting bound SOCS16 protein antibody.

Rational drug design may also be based upon structural studies of the molecular shapes of the SOCS16 protein and other effectors or analogs. See, e.g., Methods in Enzymology vols 202 and 203. Effectors may be other proteins which mediate other functions in response to protein binding, or other proteins which normally interact with the binding partner. One means for determining which sites interact with specific other proteins is a physical structure determination, e.g., x-ray crystallography or 2 dimensional NMR techniques. These will provide guidance as to which amino acid residues form molecular contact regions. For a detailed description of protein structural determination, see, e.g., Blundell and Johnson (1976) Protein Crystallography Academic Press, NY.

A purified SOCS16 protein can be coated directly onto plates for use in the aforementioned drug screening techniques. However, non-neutralizing antibodies to these proteins can be used as capture antibodies to immobilize the respective protein on the solid phase.

XI. Kits

This invention also contemplates use of SOCS16 proteins, fragments thereof, peptides, and their fusion products in a variety of diagnostic kits and methods for detecting the presence of SOCS16 protein or a SOCS16 binding partner. Typically the kit will have a compartment containing either a defined SOCS16 protein peptide or gene segment or a reagent which recognizes one or the other, e.g., binding partner fragments or antibodies.

A kit for determining the binding affinity of a test compound to a SOCS16 protein would typically comprise a test compound; a labeled compound, e.g., a binding agent or antibody having known binding affinity for the SOCS16 protein; a source of SOCS16 protein (naturally occurring or recombinant); and a means for separating bound from free labeled compound, such as a solid phase for immobilizing the SOCS16 protein. Once compounds are screened, those having suitable binding affinity to the SOCS16 protein can be evaluated in suitable biological assays, as are well known in the art, to determine whether they act as agonists or antagonists to the binding partner. The availability of recombinant SOCS16 protein polypeptides also provide well defined standards for calibrating such assays.

A preferred kit for determining the concentration of, for example, a SOCS16 protein in a sample would typically comprise a labeled compound, e.g., binding partner or antibody, having known binding affinity for the SOCS16 protein, a source of SOCS16 protein (naturally occurring or recombinant), and a means for separating the bound from free labeled compound, for example, a solid phase for immobilizing the SOCS16 protein. Compartments containing reagents, and instructions, will normally be provided.

Antibodies, including antigen binding fragments, specific for the SOCS16 protein or fragments thereof are useful in diagnostic applications to detect the presence

of elevated levels of SOCS16 protein and/or its fragments. Such diagnostic assays can employ lysates, live cells, fixed cells, immunofluorescence, cell cultures, body fluids, and further can involve the detection of antigens related to the protein in serum, or the like. Diagnostic assays may be homogeneous (without a separation step between free reagent and antigen-SOCS16 protein complex) or heterogeneous (with a separation step). Various commercial assays exist, such as radioimmunoassay (RIA), enzyme-linked immunosorbentassay (ELISA), enzyme immunoassay (EIA), enzyme-multiplied immunoassay technique (EMIT), substrate-labeled fluorescent immunoassay (SLFIA), and the like. For example, unlabeled antibodies can be employed by using a second antibody which is labeled and which recognizes the antibody to a SOCS16 protein or to a particular fragment thereof. Similar assays have also been extensively discussed in the literature. See, e.g., Harlow and Lane (1988) Antibodies: A Laboratory Manual, CSH Press, NY; Chan (ed.) (1987) Immunoassay: A Practical Guide Academic Press, Orlando, FL; Price and Newman (eds.) (1991) Principles and Practice of Immunoassay Stockton Press, NY; and Ngo (ed.) (1988) Nonisotopic Immunoassay Plenum Press, NY.

Anti-idiotypic antibodies may have similar use to diagnose presence of antibodies against a SOCS16 protein, as such may be diagnostic of various abnormal states. For example, overproduction of SOCS16 protein may result in production of various immunological or other medical reactions which may be diagnostic of abnormal physiological states, e.g., in cell growth, activation, or differentiation.

Frequently, the reagents for diagnostic assays are supplied in kits, so as to optimize the sensitivity of the assay. For the subject invention, depending upon the nature of the assay, the protocol, and the label, either labeled or unlabeled antibody or binding partner, or

labeled SOCS16 protein is provided. This is usually in conjunction with other additives, such as buffers, stabilizers, materials necessary for signal production such as substrates for enzymes, and the like. Preferably, the kit will also contain instructions for proper use and disposal of the contents after use. Typically the kit has compartments for each useful reagent. Desirably, the reagents are provided as a dry lyophilized powder, where the reagents may be reconstituted in an aqueous medium providing appropriate concentrations of reagents for performing the assay.

Many of the aforementioned constituents of the drug screening and the diagnostic assays may be used without modification, or may be modified in a variety of ways. For example, labeling may be achieved by covalently or non-covalently joining a moiety which directly or indirectly provides a detectable signal. In any of these assays, the protein, test compound, SOCS16 protein, or antibodies thereto can be labeled either directly or indirectly. Possibilities for direct labeling include label groups: radiolabels such as ^{125}I, enzymes (U.S. Pat. No. 3,645,090) such as peroxidase and alkaline phosphatase, and fluorescent labels (U.S. Pat. No. 3,940,475) capable of monitoring the change in fluorescence intensity, wavelength shift, or fluorescence polarization. Possibilities for indirect labeling include biotinylation of one constituent followed by binding to avidin coupled to one of the above label groups.

There are also numerous methods of separating the bound from the free protein, or alternatively the bound from the free test compound. The SOCS16 protein can be immobilized on various matrices followed by washing. Suitable matrices include plastic such as an ELISA plate, filters, and beads. Methods of immobilizing the SOCS16 protein to a matrix include, without limitation, direct adhesion to plastic, use of a capture antibody, chemical

coupling, and biotin-avidin. The last step in this approach involves the precipitation of protein/binding partner or antigen/antibody complex by any of several methods including those utilizing, e.g., an organic solvent such as polyethylene glycol or a salt such as ammonium sulfate. Other suitable separation techniques include, without limitation, the fluorescein antibody magnetizable particle method described in Rattle, et al. (1984) Clin. Chem. 30:1457-1461, and the double antibody magnetic particle separation as described in U.S. Pat. No. 4,659,678.

Methods for linking proteins or their fragments to the various labels have been extensively reported in the literature and do not require detailed discussion here. Many of the techniques involve the use of activated carboxyl groups either through the use of carbodiimide or active esters to form peptide bonds, the formation of thioethers by reaction of a mercapto group with an activated halogen such as chloroacetyl, or an activated olefin such as maleimide, for linkage, or the like. Fusion proteins will also find use in these applications.

Another diagnostic aspect of this invention involves use of oligonucleotide or polynucleotide sequences taken from the sequence of a SOCS16 protein. These sequences can be used as probes for detecting levels of the SOCS16 protein message in samples from natural sources, or patients suspected of having an abnormal condition, e.g., cancer or developmental problem. The preparation of both RNA and DNA nucleotide sequences, the labeling of the sequences, and the preferred size of the sequences has received ample description and discussion in the literature. Normally an oligonucleotide probe should have at least about 14 nucleotides, usually at least about 18 nucleotides, and the polynucleotide probes may be up to several kilobases. Various labels may be employed, most commonly radionuclides, particularly ^{32}P. However, other techniques may also be employed, such as

using biotin modified nucleotides for introduction into a polynucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides, fluorophores, enzymes, or the like. Alternatively, antibodies may be employed which can recognize specific duplexes, including DNA duplexes, RNA duplexes, DNA-RNA hybrid duplexes, or DNA-protein duplexes. The antibodies in turn may be labeled and the assay carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected. The use of probes to the novel anti-sense RNA may be carried out using many conventional techniques such as nucleic acid hybridization, plus and minus screening, recombinational probing, hybrid released translation (HRT), and hybrid arrested translation (HART). This also includes amplification techniques such as polymerase chain reaction (PCR).

Diagnostic kits which also test for the qualitative or quantitative presence of other markers are also contemplated. Diagnosis or prognosis may depend on the combination of multiple indications used as markers. Thus, kits may test for combinations of markers. See, e.g., Viallet, et al. (1989) Progress in Growth Factor Res. 1:89-97.

The broad scope of this invention is best understood with reference to the following examples, which are not intended to limit the invention to specific embodiments.

EXAMPLES

I. General Methods

Many of the standard methods below are described or referenced, e.g., in Maniatis, et al. (1982) Molecular Cloning, A Laboratory Manual Cold Spring Harbor

Laboratory, Cold Spring Harbor Press, NY; Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed.) Vols. 1-3, CSH Press, NY; Ausubel, et al., Biology Greene Publishing Associates, Brooklyn, NY; or Ausubel, et al. (1987 and Supplements) Current Protocols in Molecular Biology Wiley/Greene, NY; Innis, et al. (eds.) (1990) PCR Protocols: A Guide to Methods and Applications Academic Press, NY. Methods for protein purification include such methods as ammonium sulfate precipitation, column chromatography, electrophoresis, centrifugation, crystallization, and others. See, e.g., Ausubel, et al. (1987 and periodic supplements); Deutscher (1990) "Guide to Protein Purification," Methods in Enzymology vol. 182, and other volumes in this series; Coligan, et al. (1995 and supplements) Current Protocols in Protein Science John Wiley and Sons, New York, NY; P. Matsudaira (ed.) (1993) A Practical Guide to Protein and Peptide Purification for Microsequencing, Academic Press, San Diego, CA; and manufacturer's literature on use of protein purification products, e.g., Pharmacia, Piscataway, NJ, or Bio-Rad, Richmond, CA. Combination with recombinant techniques allow fusion to appropriate segments (epitope tags), e.g., to a FLAG sequence or an equivalent which can be fused, e.g., via a protease-removable sequence. See, e.g., Hochuli (1989) Chemische Industrie 12:69-70; Hochuli (1990) "Purification of Recombinant Proteins with Metal Chelate Absorbent" in Setlow (ed.) Genetic Engineering, Principle and Methods 12:87-98, Plenum Press, NY; and Crowe, et al. (1992) OIAexpress: The High Level Expression & Protein Purification System QUIAGEN, Inc., Chatsworth, CA.

Standard immunological techniques are described, e.g., in Hertzenberg, et al. (eds. 1996) Weir's Hanbook of Experimental Immunology vols 1-4, Blackwell Science; Coligan (1991) Current Protocols in Immunology Wiley/Greene, NY; and Methods in Enzymology volumes. 70, 73, 74, 84, 92, 93, 108, 116, 121, 132, 150, 162, and

163. Assays for neural cell biological activities are described, e.g., in Wouterlood (ed. 1995) Neuroscience Protocols modules 10, Elsevier; Methods in Neurosciences Academic Press; and Neuromethods Humana Press, Totowa, NJ. Methodology of developmental systems is described, e.g., in Meisami (ed.) Handbook of Human Growth and Developmental Biology CRC Press; and Chrispeels (ed.) Molecular Techniques and Approaches in Developmental Biology Interscience.

FACS analyses are described in Melamed, et al. (1990) Flow Cytometry and Sorting Wiley-Liss, Inc., New York, NY; Shapiro (1988) Practical Flow Cytometry Liss, New York, NY; and Robinson, et al. (1993) Handbook of Flow Cytometry Methods Wiley-Liss, New York, NY.

II. Isolation of full length human SOCS16 clones

Standard methods are used to isolate full length genes. A cDNA library from a human cell, preferably a STAT containing cell type. The appropriate sequence is selected, and hybridization at high stringency conditions is performed to find a full length corresponding gene. It is noted that the mouse and human protein sequences are virtually identical.

III. Isolation of primate SOCS16 clones

The full length, or appropriate fragments, of human genes are used to isolate a corresponding monkey or other primate gene. Preferably a full length coding sequence is used for hybridization. Similar source materials as indicated above are used to isolate natural genes, including genetic, polymorphic, allelic, or strain variants. Other species variants are also isolated using similar methods.

IV. Isolation of an avian SOCS16 clone

An appropriate avian source is selected as above. Similar methods are utilized to isolate a species

variant, though the level of similarity will typically be lower for avian protein as compared to a human to mouse sequence.

V. Expression; purification; characterization

Proteins of interest are immunoprecipitated and affinity purified as described above, e.g., from a natural or recombinant source.

Alternatively, with an appropriate clone from above, the coding sequence is inserted into an appropriate expression vector. This may be in a vector specifically selected for a prokaryote, yeast, insect, or higher vertebrate, e.g., mammalian expression system. Standard methods are applied to produce the gene product, preferably as a soluble secreted molecule, but will, in certain instances, also be made as an intracellular protein. Intracellular proteins typically require cell lysis to recover the protein, and insoluble inclusion bodies are a common starting material for further purificiation.

With a clone encoding a vertebrate SOCS16 protein, recombinant production means are used, although natural forms may be purified from appropriate sources. The protein product is purified by standard methods of protein purification, in certain cases, e.g., coupled with immunoaffinity methods. Immunoaffinity methods are used either as a purification step, as described above, or as a detection assay to determine the separation properties of the protein.

Preferably, the protein is secreted into the medium, and the soluble product is purified from the medium in a soluble form. Alternatively, as described above, inclusion bodies from prokaryotic expression systems are a useful source of material. Typically, the insoluble protein is solubilized from the inclusion bodies and refolded using standard methods. Purification methods are developed as described above.

The product of the purification method described above is characterized to determine many structural features. Standard physical methods are applied, e.g., amino acid analysis and protein sequencing. The resulting protein is subjected to CD spectroscopy and other spectroscopic methods, e.g., NMR, ESR, mass spectroscopy, etc. The product is characterized to determine its molecular form and size, e.g., using gel chromatography and similar techniques. Understanding of the chromatographic properties will lead to more gentle or efficient purification methods.

Prediction of glycosylation sites may be made, e.g., as reported in Hansen, et al. (1995) Biochem. J. 308:801-813.

The purified protein is also be used to identify other binding partners of SOCS16 as described, e.g., in Fields and Song (1989) Nature 340:245-246.

VI. Preparation of antibodies against vertebrate SOCS16

With protein produced, as above, animals are immunized to produce antibodies. Polyclonal antiserum is raised using non-purified antigen, though the resulting serum will exhibit higher background levels. Preferably, the antigen is purified using standard protein purification techniques, including, e.g., affinity chromatography using polyclonal serum indicated above. Presence of specific antibodies is detected using defined synthetic peptide fragments.

Polyclonal serum is raised against a purified antigen, purified as indicated above, or using, e.g., a plurality of, synthetic peptides. A series of overlapping synthetic peptides which encompass all of the full length sequence, if presented to an animal, will produce serum recognizing most linear epitopes on the protein. Such an antiserum is used to affinity purify protein, which is, in turn, used to introduce intact full

length protein into another animal to produce another antiserum preparation.

Similar techniques are used to generate induce monoclonal antibodies to either unpurified antigen, or, preferably, purified antigen.

VII. Cellular and tissue distribution

Distribution of the protein or gene products are determined, e.g., using immunohistochemistry with an antibody reagent, as produced above, by Western blotting of cell lysates, or by screening for nucleic acids encoding the respective protein. Either hybridization or PCR methods are used to detect DNA, cDNA, or message content. Histochemistry allows determination of the specific cell types within a tissue which express higher or lower levels of message or DNA. Antibody techniques are useful to quantitate protein in a biological sample, including a liquid or tissue sample. Immunoassays are developed to quantitate protein. Also FACS analysis may be used to evaluate expression in a cell population. Appropriate tissue samples or cell types are isolated and prepared for such detection. Commercial tissue blots are available, e.g., from Clontech (Mountain View, CA). Alternatively, cDNA library Southern blots can be analyzed.

VIII. STAT interference with SOCS16 proteins

Standard methods for testing the biological activity of the SOCS gene products in STAT signaling are described, e.g., in Starr, et al. (1997) Nature 387:917-921; Endo, et al. (1997) Nature 387:921-924; and Naka, et al. Nature 387:924-929. Alternatively, JAK/STATs are necessary for signal transduction. This assay is performed as described, e.g., in Ho, et al. (1995) Mol. Cell. Biol. 15:5043-5-53, and blockage with these gene products may be tested.

IX. Antagonists of SOCS function

The inhibition of SOCS function may be effected by inhibitors of the specific interaction of these gene products and their respective STAT molecules. With the information on the specificity of pairings between these SOCS and respective STAT family members, compound libraries may be screened for blockage of such interactions. Thus, inhibitory action of the SOCS may be blocked with small molecule drug candidates.

Methods of using gene therapy are described, e.g., in Goodnow (1992) "Transgenic Animals" in Roitt (ed.) Encyclopedia of Immunology, Academic Press, San Diego, pp. 1502-1504; Travis (1992) Science 256:1392-1394; Kuhn, et al. (1991) Science 254:707-710; Capecchi (1989) Science 244:1288; Robertson (1987)(ed.) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, IRL Press, Oxford; and Rosenberg (1992) J. Clinical Oncology 10:180-199. Also included is the use of antisense RNA in gene therapy to block expression of the target gene, or proper splicing of gene transcripts.

X. Comparison of various SOCS embodiments

Table 1 shows comparison of various SOCS embodiments. The "SOCSBOX protein" is a consensus of the mouse and human SOCS15 protein sequences, which are identical, but described in the filing by Johnston with Attorney Docket Number DX0761, and incorporated herein by reference. See GenBank Accession numbers U88325; U88326; U88327; U88328; AB000676; AB000677; AB000710.

Table 1: Comparison of various SOCS family members. mCIS is SEQ ID NO: 7; hSOCS1 IS SEQ ID NO: 8; mSOCS1 is SEQ ID NO: 9; hSOCS2 is SEQ ID NO: 10; hSOCS3 is SEQ ID NO: 11; mSOCS3 is SEQ ID NO: 12.

```
mCIS
hSOCS1
mSOCS1
hSOCS2
hSOCS3
mSOCS3
hSOCS14                                                MEVRVKALVHSSS
hSOCS16
mSOCS17                                 AELGEIR-----------PESAQKKLPLRKA
hSOCS18  MDKVGKMWNNLKYRCQNLFSHEGGSRNENVEMNPNRCPSVKEKSISLGEA
hSOCS19                                 ERGLETNSCSEEELSSPGRGGGGGGRLLLQ


mCIS
hSOCS1
mSOCS1
hSOCS2                          ALSPAATLTAWPADSARRGP-------------
hSOCS3
mSOCS3
hSOCS14  PSPALNGVRKDFHDLQSETTCQEQANSLKSSASHNGDLHLHLDEHVPVVI
hSOCS16
mSOCS17  EN-----------TIFITLEIVKNLFKMAENNSKNVDVRPKTSRSRSAD-
hSOCS18  APQQESSPLRENVALQLGLSPSKTFSRRNQNCAAEIPQVVEISIEKDSDS
hSOCS19  PPGPELPPVPFPLQDLVPLGRLSRGEQQQQQQQQPPPPPPPPGPLRPLAG


mCIS
hSOCS1
mSOCS1
hSOCS2   --------------------------------------------------
hSOCS3
mSOCS3
hSOCS14  G-----------------LMPQDYIQYTVPLDEGMYPLEGSRS-------
hSOCS16
mSOCS17  ---------RKD----GYVWSGKK-LSWSKKSESCSESEAKKG-------
hSOCS18  GATPGTRLARRDSYSRHAPWGGKKKHSCSTKTQSSLDTEKKFGRTRSGLQ
hSOCS19  -----------------PSRKGSFKIRLSRLFRTKSCNGGSGG-------


mCIS                                                    MVLCVQG
hSOCS1
mSOCS1
hSOCS2   -----GCTASGYPVPAARA-PAAGDQWVT--AAARDFVIR--PPGSGEKE
hSOCS3
mSOCS3
hSOCS14  -----YCLDSSSPMEVSAVPPQVGGRAFPEDESQVDQDLVVAPEIFVDQS
hSOCS16
mSOCS17  ---QLSCSSIELDLDHSCG-HRFLGRSLK--QKLQDAVGQCFPIKNCSGR
hSOCS18  RRERRYGVSSMQDMDSVSS-RAVGSRSLR--QRLQDTVGLCFPMRTYSKQ
hSOCS19  ----GDGTGKRPSGELAAS-AASLTDMGG--SAGRELDAGRKPKLTRTQS
```

Table 1 (continued):

```
mCIS      SCPLLAVEQIGRR-PLWAQSLELPGPA----------MQPLPTGA---
hSOCS1                    MVAHNQVAADN----------AVSTAAEPR---
mSOCS1                    MVARNQVAADN----------AISPAAEPR---
hSOCS2    PHPFSLCHHFGHPAGLVLGFALTSRKD--------ANPSLTPARAAT---
hSOCS3                    MVTHSKFPAAG----------MSRPLDTSL---
mSOCS3                    MVTHSKFPAAG----------MSRPLDTSL---
hSOCS14   VNGLLIGTTGVMLQSPRAGHDDVPPLS----------PLLPPMQNNQ---
hSOCS16                     MGRAELLEGK----------MSTQDPSD----
mSOCS17   HSPGLPSKRKIHISELMLDXCXFPPRSDLAFRWHFIKRHTVPMSPNS---
hSOCS18   SKPLFSNKRKIHLSELMLEKCPFPAGSDLAQKWHLIKQHTAPVSPHSTFF
hSOCS19   AFSPVSFSPLFTGETVSLVDVDISQRG----------LTSPHPPTP---

mCIS      --------------------------------------------------
hSOCS1    ----------------RRPE----------PSSSSSSS--------PAA
mSOCS1    ----------------RRSE----------PSSSSSSSS-------PAA
hSOCS2    -------------CLCRGD----------PS----------LMTLR
hSOCS3    ----------------R--------------------------------
mSOCS3    ----------------R--------------------------------
hSOCS14   ---------------IQRNFS-------------------------GLT
hSOCS16   --------------------------------------------------
mSOCS17   ----DEWVSADLSERKLRDAQLKRRNTEDDIPCFSHTNGQPCVITANSAS
hSOCS18   DTFDPSLVSTEDEEDRLRERRRLSIEEGVDPPPNAQIHTFEATAQVNPLF
hSOCS19   -------------PPPPRRSLSLLDDISGTLPTSVLVAPMGSSLQSFPLP

mCIS      -FPEEVTEETPVQAENE----------------PKVLDP----------
hSOCS1    PARPRPCPAVPAPAPGD--------------THFRTFRS----------
mSOCS1    PVRPRPCPAVPAPAPGD--------------THFRTFRS----------
hSOCS2    CLEPSGNGGEGTRSQWG--------------TAGSAEEP----------
hSOCS3    ----------------------------------LKTFSS----------
mSOCS3    ----------------------------------LKTFSS----------
hSOCS14   GTEAHVAESMRCHLNFD---------PNSAPGVARVYDSVQ---------
hSOCS16   --------------------------------------------------
mSOCS17   CTGGHITGSMMNLVTNN-SIEDSDMDSEDEIITLCTSSRKRNKPR--WEM
hSOCS18   KLGPKLAPGMTEISGDSSAIPQANCDSEEDTTTLCLQSR-RQKQRQISGD
hSOCS19   PPPPPHAPDAFPRIAPIR----------AAESLHSQPP----------

mCIS      -------------EGDLLCIAKTFSYLRES---GWYWGSITASEARQHLQ
hSOCS1    -------------HADYRRITRASALLDAC---GFYWGPLSVHGAHERLR
mSOCS1    -------------HSDYRRITRTSALLDAC---GFYWGPLSVHGAHERLR
hSOCS2    -------------SPQAARLAKALRELGQT---GWYWGSMTVNEAKEKLK
hSOCS3    -------------KSEYQLVVNAVRKLQES---GFYWSAVTGGEANLLLS
mSOCS3    -------------KSEYQLVVNAVRKLQES---GFYWSAVTGGEANLLLS
hSOCS14   -------------SSGPMVVTSLTEELKKLAKQGWYWGPITRWEAEGKLA
hSOCS16   ---------------LWSRSDGEAELLQDL---GWYHGNLTRHAAEALLLS
mSOCS17   EEEILQLEAPPKFHTQIDYVHCLVPDLLQISNNPCYWGVMDKYAAEALLE
hSOCS18   SHTHVSRQGAWKVHTQIDYIHCLVPDLLQITGNPCYWGVMDRYEAEALSE
hSOCS19   -----QHLQCPLYRPDSSSFAASLRELEKC---GWYWGPMNWEDAEMKLK
                                          *       **   .   *
```

Table 1 (continued):

```
mCIS     KMPEGTFLVRDST-HPSYLFTLSVKTTRGPTNVRIEYADSSFRLDSNCLS
hSOCS1   AEPVGTFLVRDSR-QRNCFFALSVKMASGPTSIRVHFQAGRFHLDGS-R-
mSOCS1   AEPVGTFLVRDSR-QRNCFFALSVKMASGPTSIRVHFQAGRFHLDGS-R-
hSOCS2   EAPEGTFLIRDSS-HSDYLLTISVKTSAGPTNLRIEYQDGKFRLDSIICV
hSOCS3   AEPAGTFLIRDSSDQR-HFFALSVKTQSGTKNLRIQCEGGSFSLQSDPRS
mSOCS3   AEPAGTFLIRDSSDQR-HFFTLSVKTQSGTKNLRIQCEGGSFSLQSDPRS
hSOCS14  NVPDGSFLVRDSS-DDRYLLSLSFRSHGKTLHTRIEHSNGRFSFYEQPD-
hSOCS16  NGCDGSYLLRDS-NETTGLYSLSVRAKDSVKHFHVEYTGYSFKFGFN---
mSOCS17  GKPEGTFLLRDSA-QEDYLFSVSFRRYSRSLHARIEQWNHNFSFDAHDP-
hSOCS18  GKPEGTFLLRDSA-QEDYLFSVSSAATTGSLHARIEQWNHNFSFDAHDP-
hSOCS19  GKPDGSFLVRDSS-DPRYILSLSFRSQGITHHTRMEHYRGTFSLWCHPKF
           * *.**.***         ..*          *.      *

mCIS     RP-RILAFPDVVSLVQHYVASCAADTRSDSPDPAPTPALPMSKQDAPSDS
hSOCS1   -----ESFDCLFELLEHYVAAP----------------------RRMLG
mSOCS1   -----ETFDCLFELLEHYVAAP----------------------RRMLG
hSOCS2   KS-KLKQFDSVVHLIDYYVQMCKDK---------------RTGPEAPRNG
hSOCS3   TQ-PVPRFDCVLKLVYHYMPPPGAPSFP-SPPTEPSSEVPEQPSAQPLPG
mSOCS3   TQ-PVPRFDCVLKLVHHYMPPPGTPSFS-LPPTEPSSEVPEQPPAQALPG
hSOCS14  ----VERTYSIVDLIEHSIQGLENG-----------AFCYSRSRLPGSA
hSOCS16  ------EFSSLKDFVKHFAN----------------------QP---
mSOCS17  ---CVFHSPDITGLLEHYKDPSA-----------------CMFFEPLLS
hSOCS18  ---CVFHSSTVTGLLEHYKDPSS-----------------CMFFEPLLT
hSOCS19  EDRCQSVVEFIKRAIMHSKNGK---------------FLYFLRSRVPGLP
                        .  .

mCIS     VLPIPVATAVHLKLVQPFVRRSS----ARSLQHLCRLVINRLVA---DVD
hSOCS1   ---------------APLRQRR-----VRPLQELCRQRIVATVG-RENLA
mSOCS1   ---------------APLRQRR-----VRPLQELCRQRIVAAVG-RENLA
hSOCS2   --------------TVHLYLTKPLYTSAPSLQHLCRLTINKCTG---AIW
hSOCS3   SPPRRAYYIYSGGEKIPLVLSRPLSSNVATLQHLCRKTVNGHLDSYEKVT
mSOCS3   STPKRAYYIYSGGEKIPLVLSRPLSSNVATLQHLCRKTVNGHLDSYEKVT
hSOCS14  TYP-----------VRLTNPVSRFMQVRSLQYLCRFVIRQYTR-IDLIQ
hSOCS16  ------------------------------------LIGSETG-TLMVL
mSOCS17  ---------------TPLIRTFP-----FSLQHICRTVICNCTT-YDGID
hSOCS18  ---------------ISLNRTFP-----FSLQYICRAVICRCTT-YDGID
hSOCS19  PTP-----------VQLLYPVSRFSNVKSLQHLCRFRIRQLVR-IDHIP
                                      ** .**   .        .

mCIS     CLPLPRRMADYLRQYPFQL
hSOCS1   RIPLNPVLRDYLSSFPFQI
mSOCS1   RIPLNPVLRDYLSSFPFQI
hSOCS2   GLPLPTRLKDYLEEYKFQV
hSOCS3   QLPG-P-IREFLDQYDAPL
mSOCS3   QLPG-P-IREFLDQYDAPL
hSOCS14  KLPLPNKMKDYLQEKHY
hSOCS16  KHPYPRKVEEPSIYESVRVHTAMQTGRT
mSOCS17  ALPIPSPMKLYLKEYHYKSKVRLLRIDVPEQQ
hSOCS18  GLPLPSMLQDFLKEYHYKQKVRVRWLEREPVKAK
hSOCS19  DLPLPKPLISYIRKFYYYDPQEEVYLSLKEAQLISKQKQEVEPST
          .*    .  ..
```

All references cited herein are incorporated herein by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference in its entirety for all purposes.

Many modifications and variations of this invention can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. The specific embodiments described herein are offered by way of example only, and the invention is to be limited only by the terms of the appended claims, along with the full scope of equivalents to which such claims are entitled.

WHAT IS CLAIMED IS:

1. An isolated or recombinant SOCS16 polypeptide comprising at least 17 contiguous amino acids from SEQ ID NO: 2 or 4.

2. The polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 2 or 4.

3. A fusion protein comprising the polypeptide of claim 1 or 2.

4. A binding compound which specifically binds to the polypeptide of claim 1 or 2.

5. The binding compound of claim 4 which is an antibody or antibody fragment.

6. A nucleic acid encoding the polypeptide of claim 1 or 2.

7. An expression vector comprising the nucleic acid of claim 6.

8. A host cell comprising the vector of claim 7.

9. A process for recombinatly producing a polypeptide comprising culturing the host cell of claim 8 under conditions in which the polypeptide is expressed.

SEQUENCE LISTING

SEQ ID NO: 1 is primate SOCS16 nucleic acid sequence.
SEQ ID NO: 2 is primate SOCS16 amino acid sequence.
SEQ ID NO: 3 is refined primate SOCS16 nucleic acid sequence.
SEQ ID NO: 4 is refined primate SOCS16 amino acid sequence.
SEQ ID NO: 5 is IUPAC version of primate SOCS16.
SEQ ID NO: 6 is IUPAC version of primate SOCS16.
SEQ ID NO: 7 is mouse CIS amino acid sequence.
SEQ ID NO: 8 is primate SOCS1 amino acid sequence.
SEQ ID NO: 9 is murine SOCS1 amino acid sequence.
SEQ ID NO: 10 is primate SOCS2 amino acid sequence.
SEQ ID NO: 11 is primate SOCS3 amino acid sequence.
SEQ ID NO: 12 is murine SOCS3 amino acid sequence.
SEQ ID NO: 13 is primate SOCS14 polypeptide sequence.
SEQ ID NO: 14 is rodent SOCS17 polypeptide sequence.
SEQ ID NO: 15 is primate SOCS18 polypeptide sequence.
SEQ ID NO: 16 is primate SOCS19 polypeptide sequence.

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Schering Corporation
(B) STREET: 2000 Galloping Hill Road
(C) CITY: Kenilworth
(D) STATE: New Jersey
(E) COUNTRY: USA
(F) ZIP: 07033-0530

(ii) TITLE OF INVENTION: Suppressors of Cytokine Signaling; Related Reagents

(iii) NUMBER OF SEQUENCES: 16

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: Apple Macintosh
(C) OPERATING SYSTEM: Macintosh 7.5.3
(D) SOFTWARE: Microsoft Word 6.0

(v) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US
(B) FILING DATE: 17-JUL-1998
(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/055,853
(B) FILING DATE: 15-AUG-1997

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/055,804
(B) FILING DATE: 15-AUG-1997

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/053,244

(B) FILING DATE: 18-JUL-1997

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/053,153
(B) FILING DATE: 18-JUL-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Jaye P. McLaughlin
(B) REGISTRATION NUMBER: 41,211
(C) REFERENCE/DOCKET NUMBER: DX0762K PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (908)298-5056
(B) TELEFAX: (908)298-5388

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 576 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..573

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATG GGC AGA GCA GAA CTT CTA GAA GGG AAG ATG AGC ACC CAG GAT CCC      48
Met Gly Arg Ala Glu Leu Leu Glu Gly Lys Met Ser Thr Gln Asp Pro
 1               5                  10                  15

TCA GAT CTG TGG AGC AGA TCC GAT GGA GAG GCT GAG CTG CTC CAG GAC      96
Ser Asp Leu Trp Ser Arg Ser Asp Gly Glu Ala Glu Leu Leu Gln Asp
            20                  25                  30

TTG GGG TGG TAT CAC GGC AAC CTC ACA CGC CAT GCT GCT GAA GCT CTT     144
Leu Gly Trp Tyr His Gly Asn Leu Thr Arg His Ala Ala Glu Ala Leu
        35                  40                  45

CTC CTC TCA AAT GGA TGT GAC GGC AGC TAC CTT CTG AGG GAC AGC AAT     192
Leu Leu Ser Asn Gly Cys Asp Gly Ser Tyr Leu Leu Arg Asp Ser Asn
    50                  55                  60

GAG ACC ACC GGG CTG TAC TCT CTC TCT GTG AGG GCC AAA GAT TCT GTT     240
Glu Thr Thr Gly Leu Tyr Ser Leu Ser Val Arg Ala Lys Asp Ser Val
65                  70                  75                  80

AAA CAC TTT CAT GTT GAA TAT ACT GGA TAT TCA TTT AAA TTT GGC TTT     288
Lys His Phe His Val Glu Tyr Thr Gly Tyr Ser Phe Lys Phe Gly Phe
```

```
                    85                  90                  95

AAT GAA TTC TCA TCT TTG AAG GAT TTT GTC AAG CAT TTT GCA AAT CAG      336
Asn Glu Phe Ser Ser Leu Lys Asp Phe Val Lys His Phe Ala Asn Gln
            100                 105                 110

CCT TTG ATT GGA AGC GAG ACA GGC ACT CTG ATG GTT CTA AAA CAT CCC      384
Pro Leu Ile Gly Ser Glu Thr Gly Thr Leu Met Val Leu Lys His Pro
        115                 120                 125

TAC CCA AGA AAA GTG GAA GAA CCC TCC ATT TAT GAA TCT GTC CGG GTT      432
Tyr Pro Arg Lys Val Glu Glu Pro Ser Ile Tyr Glu Ser Val Arg Val
    130                 135                 140

CAC ACA GCA ATG CAG ACA GGA AGA ACA GAA GAT GAC CTT GTG CCC ACA      480
His Thr Ala Met Gln Thr Gly Arg Thr Glu Asp Asp Leu Val Pro Thr
145                 150                 155                 160

GCA CCT TCT CTG GGC ACC AAA GAA GGT TAC CTC ACC AAA CAG GGG AGG      528
Ala Pro Ser Leu Gly Thr Lys Glu Gly Tyr Leu Thr Lys Gln Gly Arg
                165                 170                 175

CCT GGT CAA GAC CTG GAA AAC AAG ATG GTT TAC TCT GCA CAG GAA          573
Pro Gly Gln Asp Leu Glu Asn Lys Met Val Tyr Ser Ala Gln Glu
            180                 185                 190

TGA                                                                  576
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 191 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Gly Arg Ala Glu Leu Leu Glu Gly Lys Met Ser Thr Gln Asp Pro
  1               5                  10                  15

Ser Asp Leu Trp Ser Arg Ser Asp Gly Glu Ala Glu Leu Leu Gln Asp
            20                  25                  30

Leu Gly Trp Tyr His Gly Asn Leu Thr Arg His Ala Ala Glu Ala Leu
        35                  40                  45

Leu Leu Ser Asn Gly Cys Asp Gly Ser Tyr Leu Leu Arg Asp Ser Asn
    50                  55                  60

Glu Thr Thr Gly Leu Tyr Ser Leu Ser Val Arg Ala Lys Asp Ser Val
65                  70                  75                  80

Lys His Phe His Val Glu Tyr Thr Gly Tyr Ser Phe Lys Phe Gly Phe
```

85 90 95

Asn Glu Phe Ser Ser Leu Lys Asp Phe Val Lys His Phe Ala Asn Gln
100 105 110

Pro Leu Ile Gly Ser Glu Thr Gly Thr Leu Met Val Leu Lys His Pro
115 120 125

Tyr Pro Arg Lys Val Glu Glu Pro Ser Ile Tyr Glu Ser Val Arg Val
130 135 140

His Thr Ala Met Gln Thr Gly Arg Thr Glu Asp Asp Leu Val Pro Thr
145 150 155 160

Ala Pro Ser Leu Gly Thr Lys Glu Gly Tyr Leu Thr Lys Gln Gly Arg
165 170 175

Pro Gly Gln Asp Leu Glu Asn Lys Met Val Tyr Ser Ala Gln Glu
180 185 190

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1530 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 42
(D) OTHER INFORMATION: /note= "Nucleotide may be C or T at positions: 42, 50, 69, and 1147."

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 45
(D) OTHER INFORMATION: /note= "Nucleotide may be A or T at positions: 45, 49, 62, 66, 68, and 1152."

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 47
(D) OTHER INFORMATION: /note= "Nucleotide may be G or C at positions: 47, 52, and 1045."

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 48
(D) OTHER INFORMATION: /note= "Nucleotide may be A or C at positions: 48, 51, 54, 65, 67, 70, and..."

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION: 64
(D) OTHER INFORMATION: /note= "Nucleotide may be G or T at position 64."

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 75
(D) OTHER INFORMATION: /note= "Nucleotide may be A or G at positions: 75, and 534."

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1164
(D) OTHER INFORMATION: /note= "Nucleotide may be A, T, C, or G at positions: 1164, 1194, 1237, 1353, and 1408."

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 134..1027

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
ACTTTACTGG AGCTCCACCG TGGTGGCGGC CGCTCTAGAA CCAGAGCCAC CCCCGGGCTG      60

CAGGCTCACC GAGCACGAAG GTGTCAGGAG CAGCCCAGTT GTGTCTCTCT CTCTACCTCT     120

GTGAAGGGCG CGA ATG GGC AGA GCA GAA CTT CTA GAA GGG AAG ATG AGC         169
                Met Gly Arg Ala Glu Leu Leu Glu Gly Lys Met Ser
                  1               5                  10

ACC CAG GAT CCC TCA GAT CTG TGG AGC AGA TCC GAT GGA GAG GCT GAG       217
Thr Gln Asp Pro Ser Asp Leu Trp Ser Arg Ser Asp Gly Glu Ala Glu
          15                  20                  25

CTG CTC CAG GAC TTG GGG TGG TAT CAC GGC AAC CTC ACA CGC CAT GCT       265
Leu Leu Gln Asp Leu Gly Trp Tyr His Gly Asn Leu Thr Arg His Ala
     30                  35                  40

GCT GAA GCT CTT CTC CTC TCA AAT GGA TGT GAC GGC AGC TAC CTT CTG       313
Ala Glu Ala Leu Leu Leu Ser Asn Gly Cys Asp Gly Ser Tyr Leu Leu
 45                  50                  55                  60

AGG GAC AGC AAT GAG ACC ACC GGG CTG TAC TCT CTC TCT GTG AGG GCC       361
Arg Asp Ser Asn Glu Thr Thr Gly Leu Tyr Ser Leu Ser Val Arg Ala
                 65                  70                  75

AAA GAT TCT GTT AAA CAC TTT CAT GTT GAA TAT ACT GGA TAT TCA TTT       409
Lys Asp Ser Val Lys His Phe His Val Glu Tyr Thr Gly Tyr Ser Phe
             80                  85                  90

AAA TTT GGC TTT AAT GAA TTC TCA TCT TTG AAG GAT TTT GTC AAG CAT       457
Lys Phe Gly Phe Asn Glu Phe Ser Ser Leu Lys Asp Phe Val Lys His
          95                 100                 105
```

```
TTT GCA AAT CAG CCT TTG ATT GGA AGC GAG ACA GGC ACT CTG ATG GTT        505
Phe Ala Asn Gln Pro Leu Ile Gly Ser Glu Thr Gly Thr Leu Met Val
    110                 115                 120

CTA AAA CAT CCC TAC CCA AGA AAA GTG GAA GAA CCC TCC ATT TAT GAA        553
Leu Lys His Pro Tyr Pro Arg Lys Val Glu Glu Pro Ser Ile Tyr Glu
125                 130                 135                 140

TCT GTC CGG GTT CAC ACA GCA ATG CAG ACA GGA AGA ACA GAA GAT GAC        601
Ser Val Arg Val His Thr Ala Met Gln Thr Gly Arg Thr Glu Asp Asp
                145                 150                 155

CTT GTG CCC ACA GCA CCT TCT CTG GGC ACC AAA GAA GGT TAC CTC ACC        649
Leu Val Pro Thr Ala Pro Ser Leu Gly Thr Lys Glu Gly Tyr Leu Thr
            160                 165                 170

AAA CAG GGA GGC CTG GTC AAG ACC TGG AAA ACA AGA TGG TTT ACT CTG        697
Lys Gln Gly Gly Leu Val Lys Thr Trp Lys Thr Arg Trp Phe Thr Leu
        175                 180                 185

CAC AGG AAT GAA CTG AAA TAC TTC AAA GAC CAG ATG TCA CCA GAA CCA        745
His Arg Asn Glu Leu Lys Tyr Phe Lys Asp Gln Met Ser Pro Glu Pro
    190                 195                 200

ATT CGG ATC CTA GAC CTA ACA GAA TGT TCA GCT GTA CAA TTC GAT TAT        793
Ile Arg Ile Leu Asp Leu Thr Glu Cys Ser Ala Val Gln Phe Asp Tyr
205                 210                 215                 220

TCA CAA GAA AGG GTA AAC TGT TTT TGT TTG GTA TTT CCA TTC AGG ACA        841
Ser Gln Glu Arg Val Asn Cys Phe Cys Leu Val Phe Pro Phe Arg Thr
                225                 230                 235

TTT TAT CTC TGT GCA AAG ACC GGA GTA GAA GCT GAT GAG TGG ATC AAG        889
Phe Tyr Leu Cys Ala Lys Thr Gly Val Glu Ala Asp Glu Trp Ile Lys
            240                 245                 250

ATA TTA CGC TGG AAA TTG TCA CAA ATA AGA AAA CAG CTC AAC CAA GGG        937
Ile Leu Arg Trp Lys Leu Ser Gln Ile Arg Lys Gln Leu Asn Gln Gly
        255                 260                 265

GAA GCA CGA TCC GAT CTC GGT CGT TCA TCT TTA AAT AGA TCT TTC TTG        985
Glu Ala Arg Ser Asp Leu Gly Arg Ser Ser Leu Asn Arg Ser Phe Leu
    270                 275                 280

CCA AGG AAT GCT CTG GCC CAG GAA CAA GTG GAA TGT TTC CCT               1027
Pro Arg Asn Ala Leu Ala Gln Glu Gln Val Glu Cys Phe Pro
285                 290                 295

TGACGCTGTG ATCTGCACCA GCTTCAAATG AAAACCGACT AAGATTTCCT TTCCAAAACA    1087

AATCAGAACC GATGCTGATT GGGACCCATA TACCACGTTG CTGACTCACT TTGCTGCCCC    1147

TCCCATATGT TGCCATCTCC TTGAGAACAC TGAAGCAATC ACCATTCTGA TAGAAAGTGC    1207

TTAAACCACC ACTCTTAGGT CTGCTCACTC TTAGAACACA CAATGGAAGA GGAAGGGTTT    1267
```

TTGTTTTCAC TCATTGTGGT CCCCAAGCCT ATTGACACTA GTTGCCTAGA GTCCCACTGT 1327

GAGTCATGGT CAGCCTGTCT GACATCCAGG TTGTGCTATT AACCAAGAAG GAAACAGATA 1387

CTTGGAGGCT TAGATGACTT CTGCAGGATT TATATTCAGA TAGAAAACAT CAAATATTTT 1447

CAGGGGAGAG GTTTTTTTTT TTAATTTTTC CCCCTTTATA CAAAAAAAAA AAAAAAAAAA 1507

CTCGAGGGGG GCCCGTACCC AAA 1530

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 298 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Arg Ala Glu Leu Leu Glu Gly Lys Met Ser Thr Gln Asp Pro
1 5 10 15

Ser Asp Leu Trp Ser Arg Ser Asp Gly Glu Ala Glu Leu Leu Gln Asp
20 25 30

Leu Gly Trp Tyr His Gly Asn Leu Thr Arg His Ala Ala Glu Ala Leu
35 40 45

Leu Leu Ser Asn Gly Cys Asp Gly Ser Tyr Leu Leu Arg Asp Ser Asn
50 55 60

Glu Thr Thr Gly Leu Tyr Ser Leu Ser Val Arg Ala Lys Asp Ser Val
65 70 75 80

Lys His Phe His Val Glu Tyr Thr Gly Tyr Ser Phe Lys Phe Gly Phe
85 90 95

Asn Glu Phe Ser Ser Leu Lys Asp Phe Val Lys His Phe Ala Asn Gln
100 105 110

Pro Leu Ile Gly Ser Glu Thr Gly Thr Leu Met Val Leu Lys His Pro
115 120 125

Tyr Pro Arg Lys Val Glu Glu Pro Ser Ile Tyr Glu Ser Val Arg Val
130 135 140

His Thr Ala Met Gln Thr Gly Arg Thr Glu Asp Asp Leu Val Pro Thr
145 150 155 160

Ala Pro Ser Leu Gly Thr Lys Glu Gly Tyr Leu Thr Lys Gln Gly Gly
165 170 175

Leu Val Lys Thr Trp Lys Thr Arg Trp Phe Thr Leu His Arg Asn Glu

180 185 190

Leu Lys Tyr Phe Lys Asp Gln Met Ser Pro Glu Pro Ile Arg Ile Leu
195 200 205

Asp Leu Thr Glu Cys Ser Ala Val Gln Phe Asp Tyr Ser Gln Glu Arg
210 215 220

Val Asn Cys Phe Cys Leu Val Phe Pro Phe Arg Thr Phe Tyr Leu Cys
225 230 235 240

Ala Lys Thr Gly Val Glu Ala Asp Glu Trp Ile Lys Ile Leu Arg Trp
245 250 255

Lys Leu Ser Gln Ile Arg Lys Gln Leu Asn Gln Gly Glu Ala Arg Ser
260 265 270

Asp Leu Gly Arg Ser Ser Leu Asn Arg Ser Phe Leu Pro Arg Asn Ala
275 280 285

Leu Ala Gln Glu Gln Val Glu Cys Phe Pro
290 295

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1532 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ACTTTACTGG AGCTCCACCG TGGTGGCGGC CGCTCTAGAA CTAGTGGATC CCCCGGGCTG 60

CAGGAAATTC GGCACGAGGT GTCAGGGAGC AGCCCAKTTG TGTCTCTCTC TCTACCTCTG 120

TGAAGGGCGC GAATGGGCAG AGCAGAACTT CTAGAAGGGA AGATGAGCAC CCAGGATCCC 180

TCAGATCTGT GGAGCAGATC CGATGGAGAG GCTGAGCTGC TCCAGGACTT GGGGTGGTAT 240

CACGGCAACC TCACACGCCA TGCTGCTGAA GCTCTTCTCC TCTCAAATGG ATGTGACGGC 300

AGCTACCTTC TGAGGGACAG CAATGAGACC ACCGGGCTGT ACTCTCTCTC TGTGAGGGCC 360

AAAGATTCTG TTAAACACTT TCATGTTGAA TATACTGGAT ATTCATTTAA ATTTGGCTTT 420

AATGAATTCT CATCTTTGAA GGATTTTGTC AAGCATTTTG CAAATCAGCC TTTGATTGGA 480

AGCGAGACAG GCACTCTGAT GGTTCTAAAA CATCCCTACC CAAGAAAAGT GGAAGAACCC 540

TCCATTTATG AATCTGTCCG GGTTCACACA GCAATGCAGA CAGGAAGAAC AGAAGATGAC 600

CTTGTGCCCA CAGCACCTTC TCTGGGCACC AAAGAAGGTT ACCTCACCAA ACAGGGGAGG 660

CCTGGTCAAG ACCTGGAAAA CAAGATGGTT TACTCTGCAC AGGAATGAAC TGAAATACTT 720

CAAAGACCAG ATGTCACCAG AACCAATTCG GATCCTAGAC CTAACAGAAT GTTCAGCTGT 780

ACAATTYGAT TATTCACAAG AAAGGGTAAA CTGTTTTTGT TTGGTATTTC CATTCAGGAC 840

ATTTTATCTC TGTGCMAAGA CCGGAGTAGA AGCTGATGAG TGGATCAAGA TATTACGCTG 900

GAAATTGTCA CAAATAAGAA AACAGCTCAA CCAAGGGGAA GCACGATCCG ATCTCGGTCG 960

TTCATCTTTA AATAGATCTT TCTTGCCAAG GAATGCTCTG GCCCANGGAA CAAGTGGAAT 1020

GTTTCCCYTG ACGCTGTGAT CTGCASCAGC TTCAAATGAA AACCGACTAA GATTTCCTTT 1080

CCAAAACAAA TCAGAANCCG ATGCTGATTG GGACCCATAT ACCACGTTGC TGACTCACTT 1140

TGCTGCCCYT CCCWTATGTT GCCATNTCCT TGAGAACACT GAAGCAATCA CCATTNTGAT 1200

AGAAAGTGCT TAAACCACCA CTCTTAGGTC TGCTCACTNT TAGAACACAC AATGGAAGAG 1260

GAAGGGTTTT TGTTTTCACT CATTGTGGTC CCCAAGCCTA TTGACACTAG TTGCCTAGAG 1320

TCCCACTGTG AGTCATGGTC AGCCTGTCTG ACATNCAGGT TGTGCTATTA ACCAAGAAGG 1380

AAACAGATAC TTGGAGGCTT AGATGACTTN TGCAGGATTT ATATTCAGAT AGAAAACATC 1440

AAATATTTTC AGGGGAGAGG TTTTTTTTTT TAATTTTTCC CCCTTTATAC AAAAAAAAAA 1500

AAAAAAAAAC TCGAGGGGGG CCCGTACCCA AA 1532

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1530 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ACTTTACTGG AGCTCCACCG TGGTGGCGGC CGCTCTAGAA CYAGWGSMWY MSCMGGGCTG 60

CWGKMWMWYM GAGCRMGAAG GTGTCAGGAG CAGCCCAGTT GTGTCTCTCT CTCTACCTCT 120

GTGAAGGGCG CGAATGGGCA GAGCAGAACT TCTAGAAGGG AAGATGAGCA CCCAGGATCC 180

CTCAGATCTG TGGAGCAGAT CCGATGGAGA GGCTGAGCTG CTCCAGGACT TGGGGTGGTA 240

TCACGGCAAC CTCACACGCC ATGCTGCTGA AGCTCTTCTC CTCTCAAATG GATGTGACGG 300

CAGCTACCTT CTGAGGGACA GCAATGAGAC CACCGGGCTG TACTCTCTCT CTGTGAGGGC 360

CAAAGATTCT GTTAAACACT TTCATGTTGA ATATACTGGA TATTCATTTA AATTTGGCTT 420

TAATGAATTC TCATCTTTGA AGGATTTTGT CAAGCATTTT GCAAATCAGC CTTTGATTGG 480

AAGCGAGACA GGCACTCTGA TGGTTCTAAA ACATCCCTAC CCAAGAAAAG TGGRAGAACC 540

CTCCATTTAT GAATCTGTCC GGGTTCACAC AGCAATGCAG ACAGGAAGAA CAGAAGATGA 600

CCTTGTGCCC ACAGCACCTT CTCTGGGCAC CAAAGAAGGT TACCTCACCA AACAGGGAGG 660

CCTGGTCAAG ACCTGGAAAA CAAGATGGTT TACTCTGCAC AGGAATGAAC TGAAATACTT 720

CAAAGACCAG ATGTCACCAG AACCAATTCG GATCCTAGAC CTAACAGAAT GTTCAGCTGT 780

ACAATTCGAT TATTCACAAG AAAGGGTAAA CTGTTTTTGT TTGGTATTTC CATTCAGGAC 840

ATTTTATCTC TGTGCAAAGA CCGGAGTAGA AGCTGATGAG TGGATCAAGA TATTACGCTG 900

GAAATTGTCA CAAATAAGAA AACAGCTCAA CCAAGGGGAA GCACGATCCG ATCTCGGTCG 960

TTCATCTTTA AATAGATCTT TCTTGCCAAG GAATGCTCTG GCCCAGGAAC AAGTGGAATG 1020

TTTCCCTTGA CGCTGTGATC TGCASCAGCT TCAAATGAAA ACCGACTAAG ATTTCCTTTC 1080

CAAAACAAAT CAGAACCGAT GCTGATTGGG ACCCATATAC CACGTTGCTG ACTCACTTTG 1140

CTGCCCYTCC CWTATGTTGC CATNTCCTTG AGAACACTGA AGCAATCACC ATTNTGATAG 1200

AAAGTGCTTA AACCACCACT CTTAGGTCTG CTCACTNTTA GAACACACAA TGGAAGAGGA 1260

AGGGTTTTTG TTTTCACTCA TTGTGGTCCC CAAGCCTATT GACACTAGTT GCCTAGAGTC 1320

CCACTGTGAG TCATGGTCAG CCTGTCTGAC ATNCAGGTTG TGCTATTAAC CAAGAAGGAA 1380

ACAGATACTT GGAGGCTTAG ATGACTTNTG CAGGATTTAT ATTCAGATAG AAAACATCAA 1440

ATATTTTCAG GGGAGAGGTT TTTTTTTTTA ATTTTTCCCC CTTTATACAA AAAAAAAAAA 1500

AAAAAAACTC GAGGGGGGCC CGTACCCAAA 1530

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 257 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Val Leu Cys Val Gln Gly Ser Cys Pro Leu Leu Ala Val Glu Gln
1 5 10 15

Ile Gly Arg Arg Pro Leu Trp Ala Gln Ser Leu Glu Leu Pro Gly Pro
20 25 30

Ala Met Gln Pro Leu Pro Thr Gly Ala Phe Pro Glu Glu Val Thr Glu
35 40 45

Glu Thr Pro Val Gln Ala Glu Asn Glu Pro Lys Val Leu Asp Pro Glu
50 55 60

Gly Asp Leu Leu Cys Ile Ala Lys Thr Phe Ser Tyr Leu Arg Glu Ser
65 70 75 80

Gly Trp Tyr Trp Gly Ser Ile Thr Ala Ser Glu Ala Arg Gln His Leu
85 90 95

Gln Lys Met Pro Glu Gly Thr Phe Leu Val Arg Asp Ser Thr His Pro
100 105 110

Ser Tyr Leu Phe Thr Leu Ser Val Lys Thr Thr Arg Gly Pro Thr Asn
115 120 125

Val Arg Ile Glu Tyr Ala Asp Ser Ser Phe Arg Leu Asp Ser Asn Cys
130 135 140

Leu Ser Arg Pro Arg Ile Leu Ala Phe Pro Asp Val Val Ser Leu Val
145 150 155 160

Gln His Tyr Val Ala Ser Cys Ala Ala Asp Thr Arg Ser Asp Ser Pro
165 170 175

Asp Pro Ala Pro Thr Pro Ala Leu Pro Met Ser Lys Gln Asp Ala Pro
180 185 190

Ser Asp Ser Val Leu Pro Ile Pro Val Ala Thr Ala Val His Leu Lys
195 200 205

Leu Val Gln Pro Phe Val Arg Arg Ser Ser Ala Arg Ser Leu Gln His
210 215 220

Leu Cys Arg Leu Val Ile Asn Arg Leu Val Ala Asp Val Asp Cys Leu
225 230 235 240

Pro Leu Pro Arg Arg Met Ala Asp Tyr Leu Arg Gln Tyr Pro Phe Gln
245 250 255

Leu

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 211 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Val Ala His Asn Gln Val Ala Ala Asp Asn Ala Val Ser Thr Ala
1 5 10 15

Ala Glu Pro Arg Arg Arg Pro Glu Pro Ser Ser Ser Ser Ser Ser Ser
20 25 30

Pro Ala Ala Pro Ala Arg Pro Arg Pro Cys Pro Ala Val Pro Ala Pro
35 40 45

Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ala Asp Tyr
50 55 60

Arg Arg Ile Thr Arg Ala Ser Ala Leu Leu Asp Ala Cys Gly Phe Tyr
65 70 75 80

Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala Glu
85 90 95

Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys Phe
100 105 110

Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg Val
115 120 125

His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Ser Phe
130 135 140

Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg Arg
145 150 155 160

Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln Glu
165 170 175

Leu Cys Arg Gln Arg Ile Val Ala Thr Val Gly Arg Glu Asn Leu Ala
180 185 190

Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe Pro
195 200 205

Phe Gln Ile
210

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 212 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Met Val Ala Arg Asn Gln Val Ala Ala Asp Asn Ala Ile Ser Pro Ala
1               5                  10                  15

Ala Glu Pro Arg Arg Arg Ser Glu Pro Ser Ser Ser Ser Ser Ser Ser
            20                  25                  30

Ser Pro Ala Ala Pro Val Arg Pro Arg Pro Cys Pro Ala Val Pro Ala
        35                  40                  45

Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp
    50                  55                  60

Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe
65                  70                  75                  80

Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala
                85                  90                  95

Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys
            100                 105                 110

Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg
        115                 120                 125

Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Thr
    130                 135                 140

Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg
145                 150                 155                 160

Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln
                165                 170                 175

Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu
            180                 185                 190

Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe
        195                 200                 205

Pro Phe Gln Ile
    210
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 306 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Ala Leu Ser Pro Ala Ala Thr Leu Thr Ala Trp Pro Ala Asp Ser Ala
1               5                  10                  15

Arg Arg Gly Pro Gly Cys Thr Ala Ser Gly Tyr Pro Val Pro Ala Ala
            20                  25                  30

Arg Ala Pro Ala Ala Gly Asp Gln Trp Val Thr Ala Ala Ala Arg Asp
        35                  40                  45

Phe Val Ile Arg Pro Pro Gly Ser Gly Glu Lys Glu Pro His Pro Phe
    50                  55                  60

Ser Leu Cys His His Phe Gly His Pro Ala Gly Leu Val Leu Gly Phe
65                  70                  75                  80

Ala Leu Thr Ser Arg Lys Asp Ala Asn Pro Ser Leu Thr Pro Ala Arg
                85                  90                  95

Ala Ala Thr Cys Leu Cys Arg Gly Asp Pro Ser Leu Met Thr Leu Arg
            100                 105                 110

Cys Leu Glu Pro Ser Gly Asn Gly Gly Glu Gly Thr Arg Ser Gln Trp
        115                 120                 125

Gly Thr Ala Gly Ser Ala Glu Glu Pro Ser Pro Gln Ala Ala Arg Leu
    130                 135                 140

Ala Lys Ala Leu Arg Glu Leu Gly Gln Thr Gly Trp Tyr Trp Gly Ser
145                 150                 155                 160

Met Thr Val Asn Glu Ala Lys Glu Lys Leu Lys Glu Ala Pro Glu Gly
                165                 170                 175

Thr Phe Leu Ile Arg Asp Ser Ser His Ser Asp Tyr Leu Leu Thr Ile
            180                 185                 190

Ser Val Lys Thr Ser Ala Gly Pro Thr Asn Leu Arg Ile Glu Tyr Gln
        195                 200                 205

Asp Gly Lys Phe Arg Leu Asp Ser Ile Ile Cys Val Lys Ser Lys Leu
```

210 215 220

Lys Gln Phe Asp Ser Val Val His Leu Ile Asp Tyr Tyr Val Gln Met
225 230 235 240

Cys Lys Asp Lys Arg Thr Gly Pro Glu Ala Pro Arg Asn Gly Thr Val
245 250 255

His Leu Tyr Leu Thr Lys Pro Leu Tyr Thr Ser Ala Pro Ser Leu Gln
260 265 270

His Leu Cys Arg Leu Thr Ile Asn Lys Cys Thr Gly Ala Ile Trp Gly
275 280 285

Leu Pro Leu Pro Thr Arg Leu Lys Asp Tyr Leu Glu Glu Tyr Lys Phe
290 295 300

Gln Val
305

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 225 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Val Thr His Ser Lys Phe Pro Ala Ala Gly Met Ser Arg Pro Leu
1 5 10 15

Asp Thr Ser Leu Arg Leu Lys Thr Phe Ser Ser Lys Ser Glu Tyr Gln
20 25 30

Leu Val Val Asn Ala Val Arg Lys Leu Gln Glu Ser Gly Phe Tyr Trp
35 40 45

Ser Ala Val Thr Gly Gly Glu Ala Asn Leu Leu Leu Ser Ala Glu Pro
50 55 60

Ala Gly Thr Phe Leu Ile Arg Asp Ser Ser Asp Gln Arg His Phe Phe
65 70 75 80

Ala Leu Ser Val Lys Thr Gln Ser Gly Thr Lys Asn Leu Arg Ile Gln
85 90 95

Cys Glu Gly Gly Ser Phe Ser Leu Gln Ser Asp Pro Arg Ser Thr Gln
100 105 110

Pro Val Pro Arg Phe Asp Cys Val Leu Lys Leu Val Tyr His Tyr Met
115 120 125

Pro Pro Pro Gly Ala Pro Ser Phe Pro Ser Pro Pro Thr Glu Pro Ser
130 135 140

Ser Glu Val Pro Glu Gln Pro Ser Ala Gln Pro Leu Pro Gly Ser Pro
145 150 155 160

Pro Arg Arg Ala Tyr Tyr Ile Tyr Ser Gly Gly Glu Lys Ile Pro Leu
165 170 175

Val Leu Ser Arg Pro Leu Ser Ser Asn Val Ala Thr Leu Gln His Leu
180 185 190

Cys Arg Lys Thr Val Asn Gly His Leu Asp Ser Tyr Glu Lys Val Thr
195 200 205

Gln Leu Pro Gly Pro Ile Arg Glu Phe Leu Asp Gln Tyr Asp Ala Pro
210 215 220

Leu
225

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 225 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Val Thr His Ser Lys Phe Pro Ala Ala Gly Met Ser Arg Pro Leu
1 5 10 15

Asp Thr Ser Leu Arg Leu Lys Thr Phe Ser Ser Lys Ser Glu Tyr Gln
20 25 30

Leu Val Val Asn Ala Val Arg Lys Leu Gln Glu Ser Gly Phe Tyr Trp
35 40 45

Ser Ala Val Thr Gly Gly Glu Ala Asn Leu Leu Leu Ser Ala Glu Pro
50 55 60

Ala Gly Thr Phe Leu Ile Arg Asp Ser Ser Asp Gln Arg His Phe Phe
65 70 75 80

Thr Leu Ser Val Lys Thr Gln Ser Gly Thr Lys Asn Leu Arg Ile Gln
85 90 95

Cys Glu Gly Gly Ser Phe Ser Leu Gln Ser Asp Pro Arg Ser Thr Gln
100 105 110

Pro Val Pro Arg Phe Asp Cys Val Leu Lys Leu Val His His Tyr Met
115 120 125

Pro Pro Pro Gly Thr Pro Ser Phe Ser Leu Pro Pro Thr Glu Pro Ser
130 135 140

Ser Glu Val Pro Glu Gln Pro Pro Ala Gln Ala Leu Pro Gly Ser Thr
145 150 155 160

Pro Lys Arg Ala Tyr Tyr Ile Tyr Ser Gly Gly Glu Lys Ile Pro Leu
165 170 175

Val Leu Ser Arg Pro Leu Ser Ser Asn Val Ala Thr Leu Gln His Leu
180 185 190

Cys Arg Lys Thr Val Asn Gly His Leu Asp Ser Tyr Glu Lys Val Thr
195 200 205

Gln Leu Pro Gly Pro Ile Arg Glu Phe Leu Asp Gln Tyr Asp Ala Pro
210 215 220

Leu
225

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 385 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Glu Val Arg Val Lys Ala Leu Val His Ser Ser Ser Pro Ser Pro
1 5 10 15

Ala Leu Asn Gly Val Arg Lys Asp Phe His Asp Leu Gln Ser Glu Thr
20 25 30

Thr Cys Gln Glu Gln Ala Asn Ser Leu Lys Ser Ser Ala Ser His Asn
35 40 45

Gly Asp Leu His Leu His Leu Asp Glu His Val Pro Val Val Ile Gly
50 55 60

Leu Met Pro Gln Asp Tyr Ile Gln Tyr Thr Val Pro Leu Asp Glu Gly

65 70 75 80

Met Tyr Pro Leu Glu Gly Ser Arg Ser Tyr Cys Leu Asp Ser Ser Ser
85 90 95

Pro Met Glu Val Ser Ala Val Pro Pro Gln Val Gly Gly Arg Ala Phe
100 105 110

Pro Glu Asp Glu Ser Gln Val Asp Gln Asp Leu Val Val Ala Pro Glu
115 120 125

Ile Phe Val Asp Gln Ser Val Asn Gly Leu Leu Ile Gly Thr Thr Gly
130 135 140

Val Met Leu Gln Ser Pro Arg Ala Gly His Asp Asp Val Pro Pro Leu
145 150 155 160

Ser Pro Leu Leu Pro Pro Met Gln Asn Asn Gln Ile Gln Arg Asn Phe
165 170 175

Ser Gly Leu Thr Gly Thr Glu Ala His Val Ala Glu Ser Met Arg Cys
180 185 190

His Leu Asn Phe Asp Pro Asn Ser Ala Pro Gly Val Ala Arg Val Tyr
195 200 205

Asp Ser Val Gln Ser Ser Gly Pro Met Val Val Thr Ser Leu Thr Glu
210 215 220

Glu Leu Lys Lys Leu Ala Lys Gln Gly Trp Tyr Trp Gly Pro Ile Thr
225 230 235 240

Arg Trp Glu Ala Glu Gly Lys Leu Ala Asn Val Pro Asp Gly Ser Phe
245 250 255

Leu Val Arg Asp Ser Ser Asp Asp Arg Tyr Leu Leu Ser Leu Ser Phe
260 265 270

Arg Ser His Gly Lys Thr Leu His Thr Arg Ile Glu His Ser Asn Gly
275 280 285

Arg Phe Ser Phe Tyr Glu Gln Pro Asp Val Glu Arg Thr Tyr Ser Ile
290 295 300

Val Asp Leu Ile Glu His Ser Ile Gln Gly Leu Glu Asn Gly Ala Phe
305 310 315 320

Cys Tyr Ser Arg Ser Arg Leu Pro Gly Ser Ala Thr Tyr Pro Val Arg
325 330 335

Leu Thr Asn Pro Val Ser Arg Phe Met Gln Val Arg Ser Leu Gln Tyr
340 345 350

Leu Cys Arg Phe Val Ile Arg Gln Tyr Thr Arg Ile Asp Leu Ile Gln
355 360 365

Lys Leu Pro Leu Pro Asn Lys Met Lys Asp Tyr Leu Gln Glu Lys His
370 375 380

Tyr
385

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 350 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Ala Glu Leu Gly Glu Ile Arg Pro Glu Ser Ala Gln Lys Lys Leu Pro
1 5 10 15

Leu Arg Lys Ala Glu Asn Thr Ile Phe Ile Thr Leu Glu Ile Val Lys
20 25 30

Asn Leu Phe Lys Met Ala Glu Asn Asn Ser Lys Asn Val Asp Val Arg
35 40 45

Pro Lys Thr Ser Arg Ser Arg Ser Ala Asp Arg Lys Asp Gly Tyr Val
50 55 60

Trp Ser Gly Lys Lys Leu Ser Trp Ser Lys Lys Ser Glu Ser Cys Ser
65 70 75 80

Glu Ser Glu Ala Lys Lys Gly Gln Leu Ser Cys Ser Ser Ile Glu Leu
85 90 95

Asp Leu Asp His Ser Cys Gly His Arg Phe Leu Gly Arg Ser Leu Lys
100 105 110

Gln Lys Leu Gln Asp Ala Val Gly Gln Cys Phe Pro Ile Lys Asn Cys
115 120 125

Ser Gly Arg His Ser Pro Gly Leu Pro Ser Lys Arg Lys Ile His Ile
130 135 140

Ser Glu Leu Met Leu Asp Thr Cys Pro Phe Pro Pro Arg Ser Asp Leu
145 150 155 160

Ala Phe Arg Trp His Phe Ile Lys Arg His Thr Val Pro Met Ser Pro
165 170 175

Asn Ser Asp Glu Trp Val Ser Ala Asp Leu Ser Glu Arg Lys Leu Arg
180 185 190

```
Asp Ala Gln Leu Lys Arg Arg Asn Thr Glu Asp Asp Ile Pro Cys Phe
        195                 200                 205

Ser His Thr Asn Gly Gln Pro Cys Val Ile Thr Ala Asn Ser Ala Ser
    210                 215                 220

Cys Thr Gly Gly His Ile Thr Gly Ser Met Met Asn Leu Val Thr Asn
225                 230                 235                 240

Asn Ser Ile Glu Asp Ser Asp Met Asp Ser Glu Asp Glu Ile Ile Thr
                245                 250                 255

Leu Cys Thr Ser Ser Arg Lys Arg Asn Lys Pro Arg Trp Glu Met Glu
            260                 265                 270

Glu Glu Ile Leu Gln Leu Glu Ala Pro Pro Lys Phe His Thr Gln Ile
        275                 280                 285

Asp Tyr Val His Cys Leu Val Pro Asp Leu Leu Gln Ile Ser Asn Asn
    290                 295                 300

Pro Cys Tyr Trp Gly Val Met Asp Lys Tyr Ala Ala Glu Ala Leu Leu
305                 310                 315                 320

Glu Gly Lys Pro Glu Gly Thr Phe Leu Leu Arg Asp Ser Ala Gln Glu
                325                 330                 335

Asp Tyr Leu Phe Ser Val Ser Phe Arg Arg Tyr Ser Arg Ser
            340                 345                 350
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 536 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Asp Lys Val Gly Lys Met Trp Asn Asn Leu Lys Tyr Arg Cys Gln
1               5                   10                  15

Asn Leu Phe Ser His Glu Gly Gly Ser Arg Asn Glu Asn Val Glu Met
            20                  25                  30

Asn Pro Asn Arg Cys Pro Ser Val Lys Glu Lys Ser Ile Ser Leu Gly
        35                  40                  45

Glu Ala Ala Pro Gln Gln Glu Ser Ser Pro Leu Arg Glu Asn Val Ala
```

50 55 60

Leu Gln Leu Gly Leu Ser Pro Ser Lys Thr Phe Ser Arg Arg Asn Gln
65 70 75 80

Asn Cys Ala Ala Glu Ile Pro Gln Val Val Glu Ile Ser Ile Glu Lys
85 90 95

Asp Ser Asp Ser Gly Ala Thr Pro Gly Thr Arg Leu Ala Arg Arg Asp
100 105 110

Ser Tyr Ser Arg His Ala Pro Trp Gly Gly Lys Lys Lys His Ser Cys
115 120 125

Ser Thr Lys Thr Gln Ser Ser Leu Asp Thr Glu Lys Lys Phe Gly Arg
130 135 140

Thr Arg Ser Gly Leu Gln Arg Arg Glu Arg Arg Tyr Gly Val Ser Ser
145 150 155 160

Met Gln Asp Met Asp Ser Val Ser Ser Arg Ala Val Gly Ser Arg Ser
165 170 175

Leu Arg Gln Arg Leu Gln Asp Thr Val Gly Leu Cys Phe Pro Met Arg
180 185 190

Thr Tyr Ser Lys Gln Ser Lys Pro Leu Phe Ser Asn Lys Arg Lys Ile
195 200 205

His Leu Ser Glu Leu Met Leu Glu Lys Cys Pro Phe Pro Ala Gly Ser
210 215 220

Asp Leu Ala Gln Lys Trp His Leu Ile Lys Gln His Thr Ala Pro Val
225 230 235 240

Ser Pro His Ser Thr Phe Phe Asp Thr Phe Asp Pro Ser Leu Val Ser
245 250 255

Thr Glu Asp Glu Glu Asp Arg Leu Arg Glu Arg Arg Arg Leu Ser Ile
260 265 270

Glu Glu Gly Val Asp Pro Pro Pro Asn Ala Gln Ile His Thr Phe Glu
275 280 285

Ala Thr Ala Gln Val Asn Pro Leu Phe Lys Leu Gly Pro Lys Leu Ala
290 295 300

Pro Gly Met Thr Glu Ile Ser Gly Asp Ser Ser Ala Ile Pro Gln Ala
305 310 315 320

Asn Cys Asp Ser Glu Glu Asp Thr Thr Thr Leu Cys Leu Gln Ser Arg
325 330 335

Arg Gln Lys Gln Arg Gln Ile Ser Gly Asp Ser His Thr His Val Ser
340 345 350

Arg Gln Gly Ala Trp Lys Val His Thr Gln Ile Asp Tyr Ile His Cys
355 360 365

Leu Val Pro Asp Leu Leu Gln Ile Thr Gly Asn Pro Cys Tyr Trp Gly
370 375 380

Val Met Asp Arg Tyr Glu Ala Glu Ala Leu Ser Glu Gly Lys Pro Glu
385 390 395 400

Gly Thr Phe Leu Leu Arg Asp Ser Ala Gln Glu Asp Tyr Leu Phe Ser
405 410 415

Val Ser Ser Ala Ala Thr Thr Gly Ser Leu His Ala Arg Ile Glu Gln
420 425 430

Trp Asn His Asn Phe Ser Phe Asp Ala His Asp Pro Cys Val Phe His
435 440 445

Ser Ser Thr Val Thr Gly Leu Leu Glu His Tyr Lys Asp Pro Ser Ser
450 455 460

Cys Met Phe Phe Glu Pro Leu Leu Thr Ile Ser Leu Asn Arg Thr Phe
465 470 475 480

Pro Phe Ser Leu Gln Tyr Ile Cys Arg Ala Val Ile Cys Arg Cys Thr
485 490 495

Thr Tyr Asp Gly Ile Asp Gly Leu Pro Leu Pro Ser Met Leu Gln Asp
500 505 510

Phe Leu Lys Glu Tyr His Tyr Lys Gln Lys Val Arg Val Arg Trp Leu
515 520 525

Glu Arg Glu Pro Val Lys Ala Lys
530 535

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 458 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Glu Arg Gly Leu Glu Thr Asn Ser Cys Ser Glu Glu Glu Leu Ser Ser
1 5 10 15

Pro Gly Arg Gly Gly Gly Gly Gly Gly Arg Leu Leu Leu Gln Pro Pro
20 25 30

```
Gly Pro Glu Leu Pro Pro Val Pro Phe Pro Leu Gln Asp Leu Val Pro
        35                  40                  45

Leu Gly Arg Leu Ser Arg Gly Glu Gln Gln Gln Gln Gln Gln Gln Gln
    50                  55                  60

Pro Pro Pro Pro Pro Pro Pro Pro Gly Pro Leu Arg Pro Leu Ala Gly
65                  70                  75                  80

Pro Ser Arg Lys Gly Ser Phe Lys Ile Arg Leu Ser Arg Leu Phe Arg
                85                  90                  95

Thr Lys Ser Cys Asn Gly Gly Ser Gly Gly Gly Asp Gly Thr Gly Lys
            100                 105                 110

Arg Pro Ser Gly Glu Leu Ala Ala Ser Ala Ala Ser Leu Thr Asp Met
        115                 120                 125

Gly Gly Ser Ala Gly Arg Glu Leu Asp Ala Gly Arg Lys Pro Lys Leu
    130                 135                 140

Thr Arg Thr Gln Ser Ala Phe Ser Pro Val Ser Phe Ser Pro Leu Phe
145                 150                 155                 160

Thr Gly Glu Thr Val Ser Leu Val Asp Val Asp Ile Ser Gln Arg Gly
                165                 170                 175

Leu Thr Ser Pro His Pro Pro Thr Pro Pro Pro Pro Pro Arg Arg Ser
            180                 185                 190

Leu Ser Leu Leu Asp Asp Ile Ser Gly Thr Leu Pro Thr Ser Val Leu
        195                 200                 205

Val Ala Pro Met Gly Ser Ser Leu Gln Ser Phe Pro Leu Pro Pro Pro
    210                 215                 220

Pro Pro Pro His Ala Pro Asp Ala Phe Pro Arg Ile Ala Pro Ile Arg
225                 230                 235                 240

Ala Ala Glu Ser Leu His Ser Gln Pro Pro Gln His Leu Gln Cys Pro
                245                 250                 255

Leu Tyr Arg Pro Asp Ser Ser Ser Phe Ala Ala Ser Leu Arg Glu Leu
            260                 265                 270

Glu Lys Cys Gly Trp Tyr Trp Gly Pro Met Asn Trp Glu Asp Ala Glu
        275                 280                 285

Met Lys Leu Lys Gly Lys Pro Asp Gly Ser Phe Leu Val Arg Asp Ser
    290                 295                 300

Ser Asp Pro Arg Tyr Ile Leu Ser Leu Ser Phe Arg Ser Gln Gly Ile
305                 310                 315                 320

Thr His His Thr Arg Met Glu His Tyr Arg Gly Thr Phe Ser Leu Trp
```

325 330 335

Cys His Pro Lys Phe Glu Asp Arg Cys Gln Ser Val Val Glu Phe Ile
340 345 350

Lys Arg Ala Ile Met His Ser Lys Asn Gly Lys Phe Leu Tyr Phe Leu
355 360 365

Arg Ser Arg Val Pro Gly Leu Pro Pro Thr Pro Val Gln Leu Leu Tyr
370 375 380

Pro Val Ser Arg Phe Ser Asn Val Lys Ser Leu Gln His Leu Cys Arg
385 390 395 400

Phe Arg Ile Arg Gln Leu Val Arg Ile Asp His Ile Pro Asp Leu Pro
405 410 415

Leu Pro Lys Pro Leu Ile Ser Tyr Ile Arg Lys Phe Tyr Tyr Tyr Asp
420 425 430

Pro Gln Glu Glu Val Tyr Leu Ser Leu Lys Glu Ala Gln Leu Ile Ser
435 440 445

Lys Gln Lys Gln Glu Val Glu Pro Ser Thr
450 455

INTERNATIONAL SEARCH REPORT

national Application No
PCT/US 98/14545

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 C07K16/18 C12N15/62

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	R. STARR ET AL: "A family of cytokine-inducible inhibitors of signalling" NATURE., vol. 387, 26 June 1997, pages 917-921, XP002085491 LONDON GB cited in the application see the whole document ---- -/--	1-9

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 November 1998	09/12/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Le Cornec, N

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/14545

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	T.A. ENDO ET AL: "A new protein containing an SH2 domain that inhibits JAK kinases" NATURE., vol. 387, 26 June 1997, pages 921-924, XP002085492 LONDON GB cited in the application see the whole document	1-9
A	C. AUFFRAY ET AL: EMBL DATABASE ENTRY HSC2RB051, ACCESSION NUMBER F08128, 17 February 1995, XP002085493 & C.R. ACAD. SC. III, SI. VIE,, vol. 318, 1995, pages 263-272, see abstract	1,6
A	NATIONAL CANCER INSTITUTE, CANCER GENOME ANATOMY PROJECT(CGAP): EMBL DATABASE ENTRY HS1166385, ACCESSION NUMBER AA251658, 15 March 1997, XP002085495 & UNPUBLISHED, see abstract	1,6
A	L. HILLIER ET AL: "The WashU-Merck EST project" EMBL DATABASE ENTRY HS72983, ACCESSION NUMBER R14729,22 April 1995, XP002085496 & UNPUBLISHED, see abstract	1,6
P,A	WO 98 20023 A (THE WALTER AND ELIZA HALL INSTITUTE OF MEDICAL RESEARCH) 14 May 1998 see page 91 - page 102 see examples 11,17,18	1-9
P,A	D.J. HILTON ET AL: "Twenty proteins containing a C-terminal SOCS box form five structural classes" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA., vol. 95, January 1998, pages 114-119, XP002085497 WASHINGTON US cited in the application see abstract	1-9

5

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 98/14545

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9820023 A	14-05-1998	AU 4694397 A	29-05-1998

THIS PAGE BLANK (USPTO)